# Photosensing in LOV domains: a computational study of signal transduction pathways


**DISSERTATION**

Presented to the Faculty of Chemistry and Pharmacy
University of Regensburg, Germany

In partial fulfillment of the requirements
for the Degree of

**DOCTOR OF NATURAL SCIENCES**
(Dr. rer. nat.)

presented by
**IVAN STAMBOLIĆ**
from
Belgrade (Serbia)

Regensburg, 2017

The work presented has been supervised by Prof. Dr. Bernhard Dick.
Thesis has been submitted on 20. 02. 2017.
Thesis has been defended on 29. 03. 2017.

Supervisory committee:

| | |
|---|---|
| Committee chair: | Prof. Dr. Frank-Michael Matysik |
| Primary supervisor: | Prof. Dr. Bernhard Dick |
| Secondary supervisor: | Prof. Dr. Dominik Horinek |
| Committee member: | Prof. Dr. Andreas Möglich |

*To the less fortunate ones*

*This is one corner... Of one country, in one continent, on one planet that's a corner of a galaxy in a corner of the universe that is forever growing and shrinking and creating and destroying and never remaining the same. For a single millisecond. There is so, so much to see...*

*— J. Smith, DW*

# Acknowledgments

First of all, I would like to express my sincere gratitude to my supervisor Prof. Dr Bernhard Dick for his constant support, endless patience, continuous encouragement and for helping me grow as a researcher. He has been a tremendous mentor. All advice and help I received was invaluable. I would not have accomplished this work without his guidance and his unique sense of humor. I am very grateful to him for dedicating so much time training and guiding me how to become a better scientist.

I would like to thank Prof. Dr Dominik Horinek for taking over the responsibility, for offering critical discussions and for his helpful comments, feedback and guidance in advancing this work and my scientific career. He offered me great help and encouragement during my graduate career, that I am very grateful for. In addition, I would also like to thank PD Dr Stephan Bäurle for all his thoughtful comments, support and scientific discussions.

I would like to give my thanks to the current and previous members of the Dick group for generating such a great, scientific and friendly environment for work and providing me with generous help and suggestions. I cannot think of a more inspiring and exciting place to work at. I would especially like to thank Kathrin Magerl for all her countless help, comments and selfless support, for never loosing patience with me and for all her guidance and encouragement when I needed it. I would like to thank Dr. Larisa Mataranga-Popa for her support and guidance during my time in the Dick lab. I would also like to thank Dr Sergeii Donets and Dr Anton Pershin who initiated some of the projects and provided the working framework for me.

Special thanks goes to Klaus Ziereis for his continuous technical aid and IT expertise. I would like to thank Andrea Nömeier for her administrative assistance and support, which have made my graduate life worry free. Additionally, I would like to thank the current and previous members of the GRK Research Training Group 1626 for generating helpful and critical scientific environment and for promoting the scientific way.

Finally, I would like to thank my dear friends and family, especially Milena Petković, Mladen Stojanović, Lidija Živanović, my brother Miljan Stambolić and my parents Dragana Stambolić and Zoran Stambolić. I would not have made it through without them. They needed to endure all my follies over the years and that is a true feat.

# Abstract

Light-, Oxygen- and Votlage- (LOV) domains are blue-light activated photoreceptors that control various functions in plants and algae. They contain a non-covalently bound flavin mononucleotide (FMN) in their protein core. Upon blue-light illumination, the LOV domains undergo a reversible photocycle. This photocycle is spectroscopically well understood, however the structural aspects of it are still under debate. Currently available data for several LOV domains show that the light-triggered activation induces a covalent bond formation between the C4a atom of FMN and the sulphur atom of a conserved cysteine. There is no large body of work, currently available, that shows the signal propagation beyond the immediate surrounding of FMN, as this is not easily achievable with the currently available techniques. X-ray crystallography can not completely describe the signal transduction pathway, probably due to the restraint imposed by the crystal lattice. NMR and CD techniques, on the other hand, did reveal the global changes induced with the LOV domain activation, such as dissociation and loss of helical structure of the J$\alpha$ helix. This ultimately leads to the activation of the linked effector domain or of the binding to the protein partner. However, the complete pathway within the LOV domain still remains to be fully described.

A number of photoswitches have been developed that extend the currently available optogenetic tools. The LOV2 domain of *A. sativa* (AsLOV2) is often used as the photosensitive part of these constructs. One of the recent photoswitches that was shown to control the motility and growth of living cells, was the LOV2-Rac1 fusion protein. This photoactivatable Rac1 (PA-Rac1) was shown to be reversibly activated with localized blue light illuminations. The LOV domain in this construct serves as a photo-sensor, while the Rac1 domain controls downstream effectors.

Here we show the molecular mechanism of LOV domains from *R. sphaeroides* and *C. reinhardtii* as model-systems for the mechanism guiding the LOV-part of the PA-Rac1. In RsLOV we propose a fast mechanism, propagating from the FMN-pocket, through the anti-parallel $\beta$-sheets, to the A'$\alpha$ helix, and with it, the J$\alpha$ helix. The J$\alpha$ helix forms a helix-turn-helix (HTH) motif, together with the K$\alpha$ helix, so the signal is propagated from the FMN-binding pocket to the HTH-motif. As the HTH-motif forms a dimerization surface, in RsLOV homodimers, blue-light activation also affects the association properties of RsLOV. In CrLOV we were able to show the mechanism and structural aspects guiding this LOV domain. Upon introducing a single point mutation, F41Y, the mechanism changes to an electron transfer completely abolishing adduct formation. Furthermore, with additional mu-

tation of the cysteine to an alanine or a serine, the yield and protonation state of the formed FMN radicals can be tuned. To further understand this electron transfer process occurring in CrLOV-F41Y mutant, we developed a theoretical method to predict the initial donor and terminal acceptor residues in addition to predicting the complete electron transfer pathway, within a protein.

In this work, we also show the mechanism behind the Rac1 protein and its spliced variant, Rac1b, as they form the second unit of the PA-Rac1 construct. We show that upon activation of the Rac1 domain the switch 1 and switch 2 regions are found in a closed conformation, as opposed to the Rac1b domain where the same regions form a released conformation, more open for downstream-partner binding. Moreover, the 19-amino acid insertion in Rac1b is a key structural region that controls the signal propagation and opening of the binding site. We also show the mechanisms behind the protein complexes of Rac1 and Rac1b with p67$^{phox}$ and RhoGDI, and the induced changes in the interaction surfaces formed between these domains.

Finally, using the knowledge of the model-systems, we bring the mechanisms together and show the signal transduction dynamics of the LOV2-Rac1b fusion construct. We identify the important aspects of the LOV2-Rac1b structure and the structural effects that the 19-amino acid insertion cause. LOV2 domain sits on top of the 19-amino acid insertion which acts as a loading arm. This region extends and contracts thus releasing or binding the LOV2 and Rac1b units more tightly. The dynamical changes induced in the structure upon the activation of either of the two subunits that form this photoactivatable construct are also described. This includes changes in the binding regions, interaction surface as well as the magnesium and calcium ions. While magnesium is important in nucleotide binding and recognition, our results suggest that the calcium ion plays no role in a calcium-deficient environment. The LOV2 domain of the fusion construct shows a mechanism similar to the ones seen in our model systems. The J$\alpha$ helix partially looses its helical structure, showing breakage and unwinding of the geometry. We also show the complete unfolding of the small A'$\alpha$ helix. Furthermore, our results suggest coupling of the G$\beta$ and H$\beta$ sheets which corresponds to the early stages of the $\beta$-sheet tightening process, seen in FTIR experiments for isolated LOV domains. Finally, we propose that Rac1b is controlling several structural aspects, from unit binding to nucleotide hydrolysis, while LOV2 has control over the partner binding as well as the release of the binding regions of Rac1b, responsible for the downstream effector binding.

# List of abreviations

| | |
|---|---|
| AA | amino acid |
| ADP | adenosine diphosphate |
| APBS | adaptive Poisson-Boltzmann solver |
| AsLOV | *Avena sativa* LOV domain |
| ATP | adenosine triphosphate |
| BLUF | blue light using FAD |
| COX | cyclooxygenase |
| CrLOV | *Clamidomonas reinhardtii* LOV domain |
| CRY | cryptochrome |
| DADS | decay associated difference spectra |
| DNA | deoxyribonucleic acid |
| DSSP | define secondary structure of proteins |
| ET | electron transfer |
| FAD | flavin adenine dinucleotide |
| FDPB | finite difference Poisson-Boltzmann |
| FMN | flavin mononucleotide |
| FTIR | fourier transform infrared spectroscopy |
| GAP | GTPase activating protein |
| GDP | guanosine diphosphate |
| GDS | GDP dissociation stimulator |
| GEF | guanine nucleotide exchange factor |
| GLUT1 | glucose transporter 1 |
| GTP | guanosine triphosphate |
| HTH | helix-turn-helix |
| IDO | indoleamine-2,3-dioxygenase |
| IR | infrared |
| LOV | light-oxygen-voltage |
| LPBE | linearized Poisson-Boltzmann equation |
| MD | molecular dynamics |
| NADPH | nicotinamide adenine dinucleotide phosphate |
| NO | nitric oxide |
| NOS | nitric oxide synthase |
| NOX | nicotinamide adenine dinucleotide phosphate oxydase |
| PA-Rac1 | photoactivatable Rac1 |
| PAS | Per-ARNT-Sim |
| PCET | proton coupled electron transfer |

| | |
|---|---|
| PDB | protein data bank |
| Phot | phototropin |
| PME | particle-mesh-Ewald |
| PSI | photosystem I |
| PSII | photosystem II |
| RhoGDI | Rho GDP-dissociation inhibitor |
| ROS | reactive oxygen species |
| RsLOV | *Rhodobacter sphaeroides* LOV domain |
| SADS | species associated difference spectrum |
| STAS | sulphate transporter and antisigma factor antagonist |
| SVD | singular value decomposition |
| TA | transient absorption |
| TCSPC | time correlated single photon counting |
| TPR | tetratrico-peptide repeat |
| UV | ultraviolet |
| Vis | visible |
| WT | wild type |

## Acronyms for Amino Acid Types

| | | |
|---|---|---|
| Ala | A | alanine |
| Arg | R | arginine |
| Asn | N | asparagine |
| Asp | D | aspartic acid |
| Cys | C | cysteine |
| Gln | Q | glutamine |
| Glu | E | glutamic acid |
| Gly | G | glycine |
| His | H | histidine |
| Ile | I | isoleucine |
| Leu | L | leucine |
| Lys | K | lysine |
| Met | M | methionine |
| Phe | F | phenylalanine |
| Pro | P | proline |
| Ser | S | serine |
| Thr | T | threonine |
| Trp | W | tryptophan |
| Tyr | Y | tyrosine |
| Val | V | valine |

# Contents

# 1 General introduction

## 1.1 Aim of this work

There is no large body of work, currently available, that describes the signaling dynamics past the immediate surrounding of the activation trigger point within a signaling protein domain. The description of the signaling pathways within a protein domain is very hard to do, using the currently available experimental techniques. X-ray crystallography has restraints imposed by the present crystal lattice. NMR and CD techniques have an issue that they are extremely hard to solve for protein environments and are, so far, only able to describe mostly global structural changes. The scope of this thesis is the investigation of the signal transduction dynamics of protein-based photoreceptors, photoswitches and optogenetic tools. Using molecular dynamics simulations we describe the signaling transduction within a protein domain, on a molecular level. We first show the early mechanisms of the RsLOV and CrLOV domains. In order to ascertain the electron transfer pathway as well as the initial donor and the terminal acceptor residues of the electron transfer process within LOV domains, we have developed a novel method that is able to answer these question. We have also simulated Rac1 and Rac1b proteins, together with their complexes with $p67^{phox}$ and RhoGDI, in an effort to understand the mechanisms guiding these two spliced variants. Studying these model-systems, RsLOV and CrLOV on one hand, and the Rac1 and Rac1b on the other hand, allowed us to apply the knowledge gained of their mechanisms onto the mechanism of the LOV2-Rac1b photoswitch. As a result of these simulations, we observed the early signal transduction dynamics within the studied photoswitches and photoreceptor constructs.

## 1.2 Outline

This thesis is divided into eight chapters. Chapters are mainly organized into four sections: the *introduction section* which gives the reader the required background information; the *results section* where all the results are presented without providing interpretations, as to give the reader an opportunity to draw their own conclusions from the presented results; the *discussion section* is the main section where the data interpretation is given; and finally the *conclusions section* is where the conclusions of the particular topics are discussed, based on the previously described interpretations.

In detail, the thesis is composed of the following chapters:

**Chapter 1** includes this outline together with the description of the aim of the work described in this thesis as well as the list of publications published as part of this thesis.

**Chapter 2** contains background information, including the general theories and explanations, required for easier understanding of the topics and results discussed later on in the thesis.

**Chapter 3** describes the molecular mechanism that drives the signaling of LOV domain in *R. sphaeroides* upon activation as well as the information exchange between the two subunits forming the RsLOV dimer.

**Chapter 4** shows the effects of single-point mutation on the reaction type in LOV domain from *C. reinhardtii*. This chapter, in part, describes the results obtained employing experimental techniques performed by Kathrin Magerl.

**Chapter 5** includes the description of the novel method for predicting electron transfer pathways in proteins. Additionally, the chapter describes the algorithms for predicting the initial donor and the terminal acceptor residues of an electron transfer process.

**Chapter 6** contains the molecular mechanism and behavior of Rac1 and its spliced variant, Rac1b. In this chapter, the mechanisms of their $p67^{phox}$, RhoGDI and LOV2 constructs are also described.

**Chapter 7** gives information on the methods used to generate the results presented in the main body of this thesis.

**Chapter 8** includes the summary of the presented results and conclusions.

## 1.3 Publications and presentations

**Parts of this thesis have been submitted for publication:**

- K. Magerl, I. Stambolic and B. Dick, *Switching from adduct formation to electron transfer in a light-oxygen-voltage domain containing the reactive cysteine*, Phys. Chem. Chem. Phys, (2017)

**Parts of this thesis constitute manuscripts in preparation:**

- I. Stambolic, D. Horinek and B. Dick, *Mechanism of early signal transduction of the short LOV photosensor from Rhodobacter Sphaeroides*, (2017)
- I. Stambolic, D. Horinek and B. Dick, *Early signaling mechanisms of Rac1, Rac1b and their constructs with RhoGDI, p67$^{phox}$ and LOV2 domain*, (2017)
- I. Stambolic, D. Horinek and B. Dick, *Predicting complete electron transfer pathways in proteins*, (2017)

**Additionally, I have contributed to the following report:**

- E. Peter, B. Dick, I. Stambolic and S. A. Baeurle, *Exploring the multiscale signaling behavior of phototropin1 from Chlamydomonas reinhardtii using a full-residue space kinetic Monte Carlo molecular dynamics technique*, Proteins: Structure, Function and Bioinformatics, **82**, 2018 (2014)

**Parts of this thesis have been presented both orally and in a poster form at over 15 international conferences.**

# 2 Background

## 2.1 Blue light sensing

Living organisms rely on the ability of their cells to sense and respond to extracellular stimuli in order to grow and survive. Single-cell organisms need to sense the nutrients present in the extracellular environment. Cells of multicellular organisms need to sense their neighbors in the surrounding tissues. In both cases, the cells need to adjust their metabolism as a response to this external stimuli. This adjustment is done through a series of signal transduction pathways. Each pathway is formed by a series of proteins that interact with each other in order to pass the information downstream. [1] Signal transduction also includes the stress propagation within each of the protein forming the pathway. This process of signal transduction within a protein is called the early signaling. Deregulation, disruption or alteration of any part of a signaling pathway can lead to cell death or disease. Signaling cascades start with the protein that is directly responding to the external stimuli. Activation of this receptor protein induces a series of structural changes that ultimately lead to its modified reactivity and interactions with downstream protein partners. [2] The basic mechanism of a signaling pathway and the corresponding stimuli response is shown in figure 2.1.1.

Light is one of the essential external stimuli that cells have evolved to sense and respond to. Photodynamic action regulates a vast number of cellular processes, such as cell adhesion, phototropism and circadian rhythm [3–5] in many organisms, from bacteria to plants. There are two ways a receptor senses the external stimuli. One is through small molecules ligand binding while the second way is with post-translational modifications. Here, we will discuss sensing through ligand binding, mostly by describing the light-oxygen-voltage sensing proteins. Specific responses listed above, cell adhesion, phototropism and circadian rhythm, are controlled by a common receptor family sensitive to light stimuli. These are the Light-Oxygen-Voltage- (LOV) sensitive domains. [6] LOV domains harbor a non-covalently bound flavin chromophore that senses the blue light (400-470 nm range). Upon light activation, it undergoes a reversible photocycle. [7] The first step in the light response of LOV domains is adduct formation, between the C4a atom of the flavin cofactor and the sulphur atom of the strictly conserved cysteine.

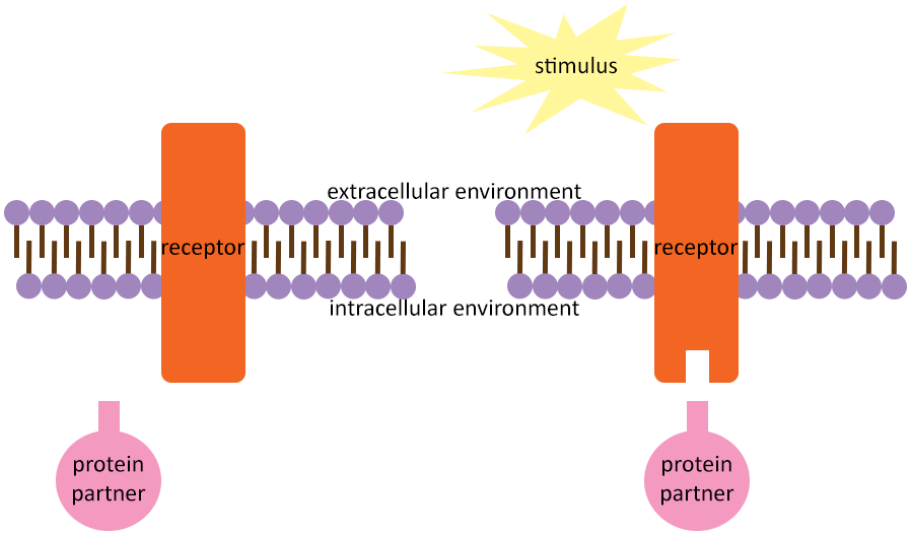


**Figure 2.1.1** Basic mechanism of the initiation of the cell signaling cascade. Upon activation by the external stimulus, receptor proteins undergo a conformational change that modifies its reactivity or partner binding properties.

### 2.1.1 Diversity of LOV domains

LOV domains are usually coupled to other domains, called effector domains. LOV domains sense the blue light and pass the information to the effector domain they are linked to. Subsequently, the effector domain then controls certain reactions in biological systems as a response to light, even though the effector domains can not sense the light themselves. LOV domains are coupled to other sensory domains, such as STAS domains, Ser/Thr kinase domains, F-box proteins and many others, depending on the organism and the particular construct in question. [8] Effector domains govern the function of these photoreceptors. Additionally, there are LOV domains that are not coupled to any effector domain, but instead they bind to protein partners in response to light. They are called short LOV domains, [9] because of their short primary sequence when compared to the sequence of the complete photoreceptor-effector construct.

### 2.1.2 LOV domain photocycle

The photocycle of a LOV domain starts from the dark-adapted state. Upon blue-light irradiation, the flavin is excited to the excited singlet state, which undergoes intersystem crossing to the excited triplet state. This reacts with a nearby cysteine residue resulting in a covalent bond between flavin C4a and cysteine sulfur atom. This photoadduct is most likely formed between a $FMNH^{\bullet}$ and $^{\bullet}SCys$ radical pair. It is the signaling state of a LOV domain. The radical mechanism is not the only mechanism proposed, [4, 7, 10] it is however, the most likely one. [11–13] In addition to this radical combination, FMN-N5 is protonated.

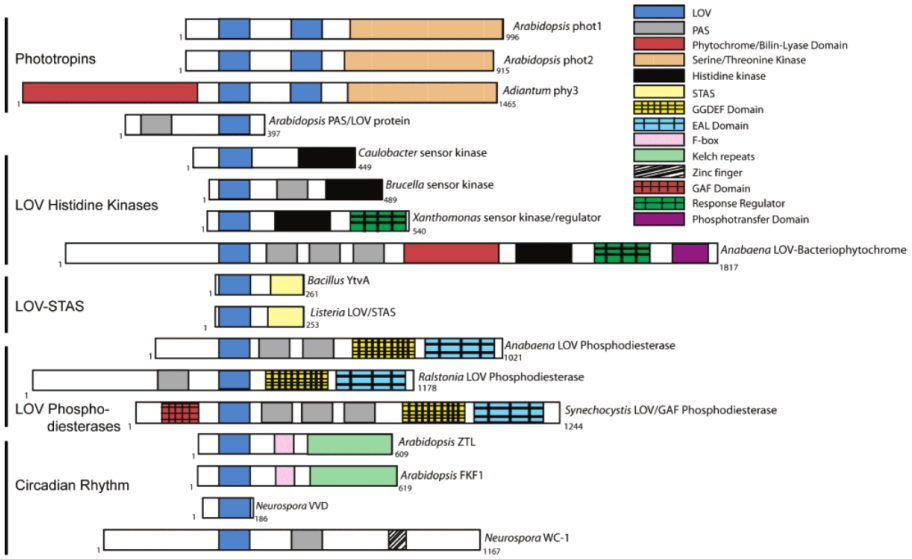


**Figure 2.1.2** Alignment of cartoon representations of a few LOV domain-containing proteins, showing a great diversity of LOV domains. Reprinted with permission from [8]. Copyright 2013 American Chemical Society.

### 2.1.3 The structure and signaling mechanism of LOV domains

There are numerous crystal structures determined for LOV domains. They show a high degree of similarity of the LOV core, because they share the same PAS (Per-ARNT-Sim) domain. Cartoon alignments of different LOV domain sequences are shown in figure 2.1.2, revealing a great diversity of LOV domains. The shared PAS core is made of antiparallel $\beta$-sheets with several $\alpha$-helices flanking the core. [14] The solvent-exposed surface, formed by the $\beta$-sheet, is identified as an interaction surface for protein-protein binding. [15]

The flanking helices vary greatly depending on the particular LOV domain. They have been shown to play a crucial role in the control of the early signaling. The C-terminal helix, named J$\alpha$ helix, associates with the $\beta$-sheet surface in the dark, while it partially looses its helical structure in the light state and detaches from the protein core. [16] This conformational change influences the linked effector domain or the protein binding surface. At the N-terminal end of the protein, there is an A'$\alpha$ helix. This helix has been shown to fully unfold upon activation and to be a key region in transducing the signal from the flavin cofactor in the protein core to the flanking helices, which ultimately pass the signal to the effector domain or the effector region. [17]

## 2.2 Electron transfer

The transport of electrons through environments in biological systems is one of the key steps in practically all processes in cells. [18] An electron transfer (ET) chain is a series of individual electron transfer steps over the charge carriers in a sequence. An electron originates from a specific site in a protein, either through excitation of a residue or by entering the protein from an outside electron donor. Similarly, it exits the protein at a different site, in most cases the substrate binding site. The protein environment controls the electron transfer pathway and specific intermediate charge carriers. Two most important features of the protein environment that guide the electron transfer process are the distances between each individual donor and acceptor as well as the local charge distribution. [19]

### 2.2.1 Models of electron transfer

Since the pioneering work by Marcus, [20, 21] several methods have been developed to treat the ET process in proteins. Two methods in particular will be described here, the Packing density model and the Pathway tunneling model. These two models are presented here as the etFIND software described later in the thesis uses these two models to determine the pathway and the two points of ET.

#### 2.2.1.1 Packing density model

The packing density or the square barrier model was developed by Hopfield in 1974. [22] This simple model approximates the tunneling medium with a square barrier. The barrier has an effective height, as shown in figure 2.2.1a. In this model, the coupling matrix element between the initial and final wave functions decays exponentially with the distance. This model is able to provide a good description of the ET in proteins. The model has been expanded for protein environments by Moser and Dutton [19, 23], where in the expanded model the rate of ET is given as equation 2.1.

$$\log k_{ET} = 13 - \frac{\beta}{2.303}(R - R_0) - \gamma\frac{(\Delta G^0 + \lambda)^2}{\lambda} \tag{2.1}$$

where $k_{ET}$ is the electron transfer rate in units of $s^{-1}$, $R$ (in Å) is the edge-to-edge electron transfer distance, $R_0$ is the van der Waals contact distance and is set to 3.6 Å, $\Delta G^0$ (in eV) is the Gibbs free energy change and $\lambda$ (in eV) is the reorganizational energy. The parameter $\gamma$ includes room temperature constants and the coefficient $\beta$ is given as a dependency of $\rho$, the packing density of the medium, which further depends on the barrier height.

$$\beta = ln10(1.2 - 0.8\rho) \tag{2.2}$$

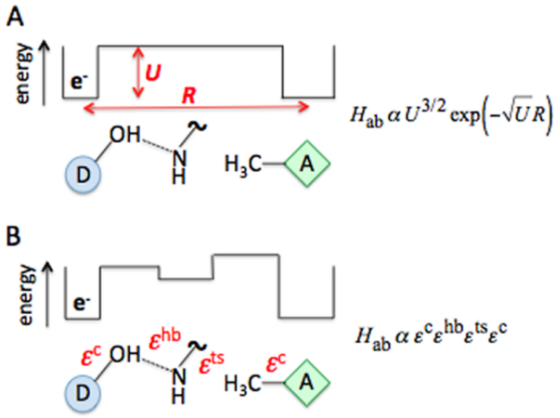


**Figure 2.2.1** Electronic coupling matrix element for ET in biological systems. Schematic representation of a) the packing density model and b) the pathway tunneling model. Adapted with permission from [18]. Copyright 2015 American Chemical Society.

### 2.2.1.2 Pathway tunneling model

The pathway model [24, 25] describes the structure of the tunneling medium on an atomistic level. Tunneling in this model takes into account the intermediate charge carriers, connecting the donor with the acceptor. This is shown in figure 2.2.1b. Electronic coupling of a given path, connecting the donor with the acceptor, is given as a product of the closest contact term, $H_{ab}$, and the attenuation factor. This attenuation factor is a product of three factors that describe the decay of consecutive tunneling through covalent bonds ($\epsilon_i^c$), hydrogen bonds ($\epsilon_j^{hb}$) and space ($\epsilon_k^{ts}$). The distance $r$ is defined as a distance between heavy atoms. This is shown in equations 2.3 to 2.6.

$$\epsilon = \prod_i \epsilon_i^c \prod_j \epsilon_j^{hb} \prod_k \epsilon_k^{ts} \tag{2.3}$$

$$\epsilon_i^c = 0.6 \tag{2.4}$$

$$\epsilon_j^{hb} = 0.36\, exp\,[-1.7(r/\text{Å} - 2.8)] \tag{2.5}$$

$$\epsilon_k^{ts} = 0.6\, exp\,[-1.7(r/\text{Å} - 1.4)] \tag{2.6}$$

The pathway model is an empirical method approximating the superexchange model, which assumes that the electron tunneling between the two ET points is controlled by the unoccupied orbitals of the atoms in-between the donor and acceptor pair.

As there is a vast number of pathways connecting donor and acceptor, this number has to be trimmed down. Usually, this is accomplished by excluding certain paths, based on geometrical parameters, such as the distance from the two points. Another

way is to use efficient search algorithms to calculate this factor for a vast number of paths. A later step includes sorting of the calculated paths as the pathway with the largest $\epsilon$ is the preferred pathway.

#### 2.2.1.3 Comparison of the two models

The extended packing density model gives good estimates of the direct tunneling between donor and acceptor. This model assumes that the environment between the donor and acceptor can be correctly described as an average medium and it does not take into account the individual charge carriers between the two ET points. The question if proteins have evolved to have a preferred pathway or it is only a matter of donor-acceptor distance has been widely debated in the literature. [26–28] However, next to this coarse graining of the protein environment, the issue with this model is that it can only describe the tunneling between the initial and final points of ET and it can not describe the individual steps, in-between these two ET points. The pathway model, on the other hand, successfully captures the atomistic structure of the protein, unlike the density packing model, but it is fully empirical.

Moreover, the two models use a different definition of the distance. In the density packing model, the distance is defined as edge-to-edge, meaning two closest atoms of the donor-acceptor residue pair. In contrast, the pathway model uses center-to-center or metal-to-metal distance of the donor-acceptor pair in its equations. The density packing model is strongly supported by the fact that when using edge-to-edge distances, ET reactions in biological systems fall within an exponential decay line. [29] However, when metal-to-metal distances are used the pathways model gives a good explanation of the resulting ET rates. [26] Long-range ET processes in biological systems are almost solely governed by the distance, but for more quantitative understanding of these reactions the structure of the protein on an atomistic scale needs to be taken into account.

## 2.3 Cancer

According to the World Health Organization [30] cancer is a collective term describing a large group of diseases that can affect any part of the body. Malignant tumors or cancers, compared to benign tumors, have the ability to infiltrate surrounding tissue which ultimately leads to the destruction of the perforated tissue. Malignant tumors also have the ability to spread to lymph nodes and distant organs via blood vessels and lymphatic pathways. This is known as metastasis.

### 2.3.1 Classification of cancer

There are over a hundred types of cancer. The classification of cancer depends on the criteria selected for classification. Classically it is based on morphological

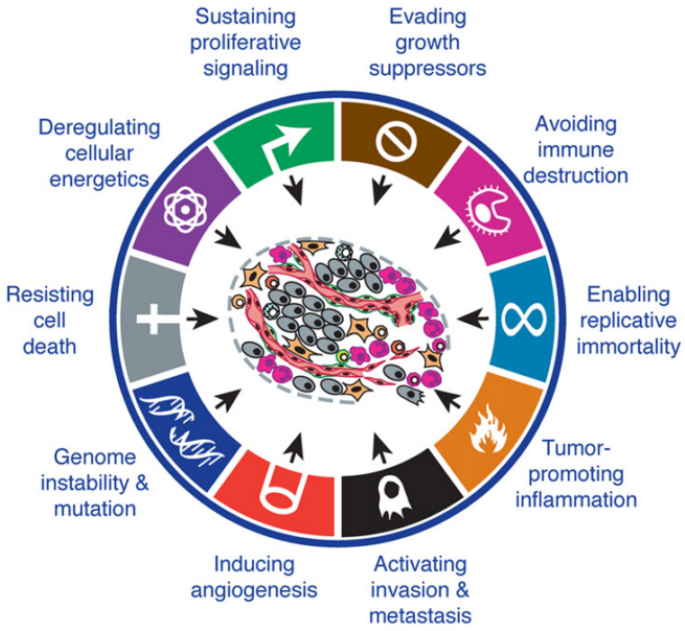


**Figure 2.3.1** Schematic representation of the hallmarks of cancer acquired during development of tumors, together with emerging hallmarks and enabling characteristics. Reprinted from [35], with permission from Elsevier.

appearance [31] or based on the cells of origin. [32] There have been other approaches [33, 34] that classify tumors based on different molecular markers. Additionally, tumors with the same histopathological appearance can have very different clinical progressions. [31] Each class of tumors have subclasses and in most cases these subclasses are distinct enough to lead to different clinical outcomes.

### 2.3.2 Development of cancer

Cancer development is a multistep process during which cancer develops several functional properties. They ensure survival, growth and dissemination of the tumor. These features are known as hallmarks of cancer [35]. Amongst others, they include sustaining proliferative signaling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis and activating invasion and metastasis as shown in figure 2.3.1.

Tumorigenesis is caused by genome instability and accumulation of mutations. The accumulation of spontaneous mutations is very slow due to efficient DNA-repair mechanisms. Tumor cells develop numerous mutations, which results in deactivation of these DNA repair mechanisms. [36] The alteration of a single gene is not enough to form cancer. Instead, subsequent mutations, enabling malignancy and invasion, are necessary for tumor progression. [37]

The most important feature of tumor cells is the sustained proliferation signaling. In cells, growth and division are normally tightly regulated by the release of growth

factors in order to maintain normal tissue function. In cancer cells this signaling is not regulated as in normal cells. The overproduction of growth factors ultimately leads to cell proliferation. This upregulation of growth factors makes the cancer cells hyperresponsive. Mutations of growth factor receptors even allow for complete independence of cancer cells from growth factors. [38] Sustained proliferation signaling can be achieved through various pathways. Prominent examples are the compromised Ras GTPase activity, [39] constitutive activation of mitogen-activated protein kinase pathway, [40] hyperactivation of AKT signal transducers [41] and loss-of-function mutations in tumor suppressor PTEN. [42]

Cancer cells have the ability to bypass apoptosis, the programmed cell death. Apoptosis is regulated through apoptotic triggering proteins BAX and BAK, BCL2-accociated X protein and BCL2 antagonist/killer, respectively. Additionally, these proteins are inhibited by BCL2 anti-apoptotic protein. Cancer cells have developed mechanism to alter the apoptosis or to completely evade it. [43] Additionally, cancer cells have developed the replicative immortality trait. It enables unlimited proliferation. [44] Activating invasion and metastasis is the final step in cancer progression. It includes dissemination of tumor cells to distant sites, through lymph channels and blood vessels. [45] A key protein involved in cell-to-cell adhesion is E-cadherin. This protein is regulated on many levels, including the tight regulation by Rac1. The downregulation and inactivation of E-cadherin is usually observed in cancer cells, [46] and can be linked with the deregulation of Rac1.

# 3 Mechanism of early signal transduction of the short LOV photosensor from *Rhodobacter sphaeroides*

## 3.1 Introduction

Flavin-binding photoreceptors are found in a variety of organisms where they control numerous functions, [47] depending on the photoreceptor family. They can be categorized [48] into three distinct groups: a) Light-, Oxygen- and Voltage- (LOV) sensing domains; b) Blue-Light Using FAD (BLUF) domains and c) Cryptochromes (CRY).

LOV domains are blue-light photoreceptors that harbor a non-covalently bound flavin mononucleotide (FMN) or in some cases flavin adenine dinucleotide (FAD) chromophore. [49] The signaling state in LOV domains is generated via blue-light induced formation of a covalent bond between the C4a atom of FMN and the sulfur atom of an adjacent cystein. [50]. In the dark, the FMN-cysteinyl adduct breaks slowly, on a time-scale from minutes to hours, reverting the LOV domain to the ground state, and concluding the photocycle. [49] This photocycle of LOV domains is shown in figure 3.1.1 panel a).

The majority of LOV domains are apart of larger multi-domain constructs, where the light-sensing of LOV domains acts as triggers or activators of linked effector domains. The LOV domain biological function in plants has been well established, however their regulatory function in bacteria is still poorly understood. [51]

In addition to multi-domain constructs, many LOV domains exist in fungi and bacteria that do not contain a coupled effector domain. They are called short LOV domains, [52] due to their short sequences when compared to the sequences of multi-domain LOV constructs. Several short LOV domains have been functionally characterized. Interestingly, short LOV domains exhibit the complete function of LOV domains, most likely through binding to other protein partners. [49] Short LOV domains contain C- and N-terminal helical extensions that are involved in signaling. [52, 53] In multi-domain LOV constructs these helical extensions link the effector domains to the LOV domain and often act as triggers or spacer providing

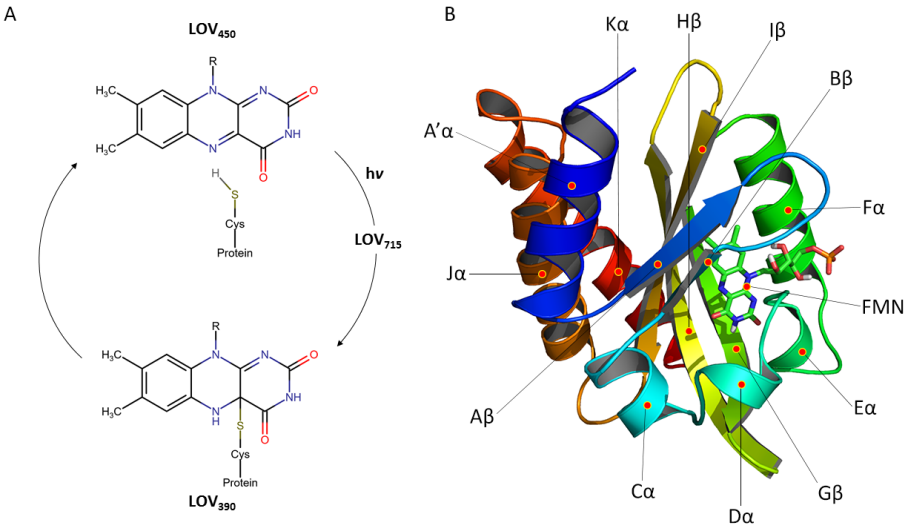


**Figure 3.1.1** (A) Simplified photocycle of a LOV domain. (B) Structure of *R. sphaeroides* LOV domain in the dark state, with all of the RsLOV structural elements labeled.

conformational flexibility. [54] A light-triggered monomer-dimer equilibrium can be a mode of action of short LOV domains where either the monomer or the dimer selectively interacts with downstream signaling partners. [17, 52, 55] This signaling control through partner binding can be mediated through the A'$\alpha$- and J$\alpha$-helices. [17]

The LOV domain photocycle starts from the dark-adapted state, $LOV_{450}$, with the non-covalently bound FMN. Upon irradiation FMN gets excited to a singlet state which then undergoes intersystem crossing to the lowest excited triplet state. From the triplet state, FMN decays into the photo-adduct, $LOV_{390}$. This photo-adduct state is most likely preceded by a transient $FMNH^{\bullet}$ - $^{\bullet}SCys$ radical pair, and represents the signaling state of the LOV domain. While the radical mechanism is not the only mechanism proposed in the literature, [4, 7, 10] it is, however, the most likely [11–13] mechanism for adduct formation in LOV domains. Following this biradical combination, which is coupled with the protonation of FMN-N5 atom, is the formation of the covalent bond between C4a atom of FMN and the sulfur atom of the cysteine residue in $LOV_{390}$ state. In most LOV domains the $LOV_{390}$, which is the signaling state *in vivo*, thermally recovers to the dark-adapted state, the $LOV_{450}$ state. [56]

One of the first and crucial steps in signal propagation to the LOV domain surface is the sidechain flip of the conserved glutamine residue that lies on the I$\beta$ strand. It has been shown multiple times [16, 57–62] that this is a critical step in the mechanism propagation. In order to maintain a hydrogen bond with the protonated N5 atom of FMN, the amide of the conserved glutamine residue must flip.

*Rhodobacter sphaeroides* is a photosynthetic purple bacteria that contains several light-sensing proteins, a short LOV domain (RsLOV), the FAD photoreceptor AppA

and three CRY-like domains. [52] RsLOV binds FMN as a chromophore. While having distinct structural characteristics, it carries some similarities to Vivid (*N. crassa*) and YtvA (*B. subtilis*). The RsLOV core is mostly conserved with respect to Vivid and YtvA LOV domains, and the highest degree of differences lie on the flanking helices. Most notably, the domain swapping N-terminal region of Vivid is lacking in RsLOV, but the receptor pocket for it is conserved. [52] Additionally, the C-terminal extension is present in RsLOV and YtvA but is lacking in Vivid. While RsLOV lacks an effector domain, the two C-terminal helices linked through a salt bridge, the J$\alpha$ and K$\alpha$ helices form a helix-turn-helix (HTH) motif, which has been shown previously to be involved in DNA and protein binding. [63, 64] The structure of the RsLOV dark state domain is shown in figure 3.1.1 panel b), with all of the structural elements labeled. In RsLOV the HTH motif is used as an unusual dimerization region, forming a tight four-helical bundle, flanked by the A'$\alpha$ helices. RsLOV is a homodimer in the dark, which dissociates upon light irradiation. [52] However, the crystal structure reveals that the dimers are made of two distinct subunits.

Interaction partners and targets of RsLOV are still unknown, but RsLOV has been linked with photosynthetic gene expression, carbohydrate metabolism, chemotaxis, response to photooxidative stress and light-controlled gene expression. [52, 53] However, the mechanism behind the RsLOV control of these processes and the light-driven signal propagation is still poorly understood.

Here, we address the issue of the short LOV photoreceptor early signaling mechanism triggered by the blue-light irradiation, in *R. sphaeroides*. We present the results of molecular dynamics simulations of both dimers and monomers of RsLOV wild type in the dark and light states. We further show the results from the simulations of different RsLOV constructs in an effort to elucidate the early signal propagation pathway in this short LOV domain. These results lead us to a proposal for the early signaling mechanism of RsLOV as well as the mechanism of unit dissociation upon blue-light activation.

## 3.2 Results

We performed non-invasive MD simulations of dark- and light-state monomers as well as the second dark domain after light-dark simulation has completed. Furthermore, we performed simulations of dark-dark, light-dark, light-light and symmetrical dark-dark dimers.

### 3.2.1 Simulations of dark-dark and light-dark dimers

Formation of the FMN-Cys covalent bond causes a perturbation in the FMN-binding pocket. The residues surrounding the FMN show significant changes in the light

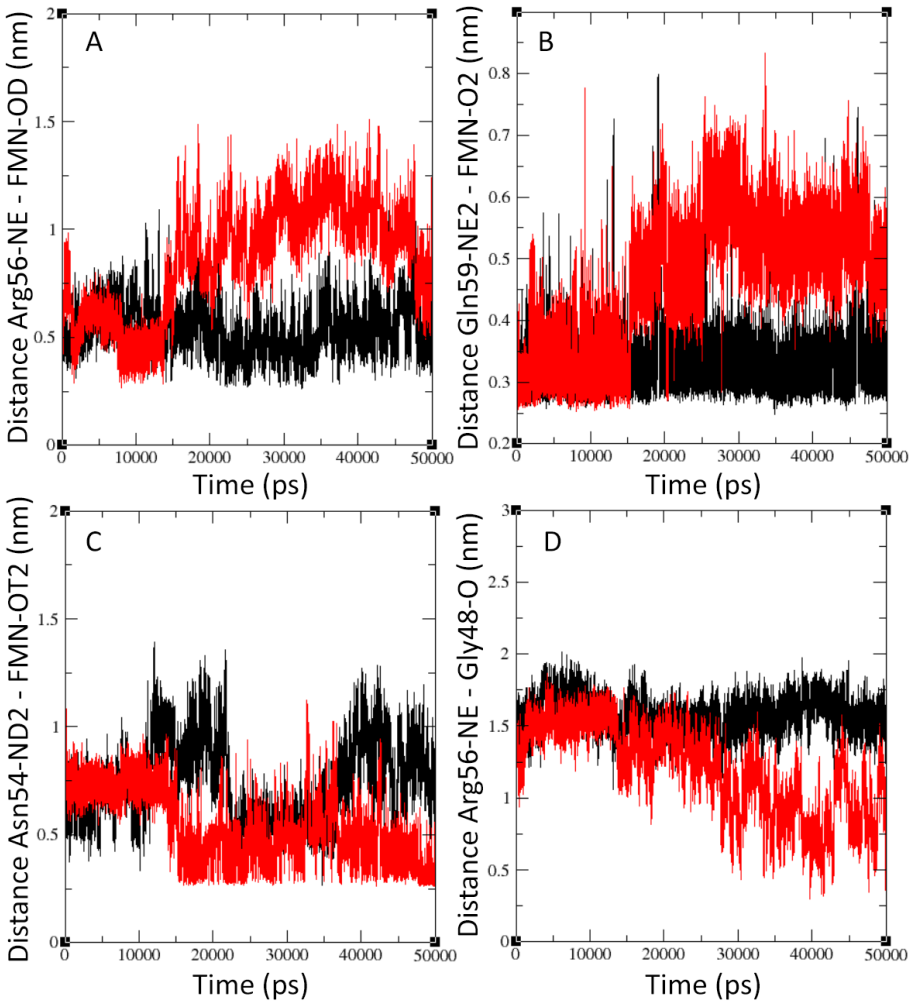


**Figure 3.2.1** Distortion of the FMN-binding pocket. Characteristic distances between (A) Arg56 and FMN, (B) Gln59 and FMN, (C) Asn54 and FMN and (D) Arg56 and Gly48 residues. The distances of the dark-dark simulation are shown in black in all graphs, while the light-dark distances are represented with red color.

state. Figure 3.2.1 shows a comparison of the behaviors of certain residues in unit A in the dark-dark and light-dark simulations. Figure 3.2.1 panels a) and b) show the separation of the E$\alpha$ helix and the FMN cofactor in the light state. In panel c) of figure 3.2.1 the closing of the D$\alpha$-E$\alpha$ loop with the FMN residue in the light state can be seen, as well as the formation of a hydrogen bond between the two regions. Figure 3.2.1 panel d) shows a hydrogen bond formation between Arg56 and Gly48, which lie on the E$\alpha$ and D$\alpha$ helices, respectively.

Upon blue-light activation, the signal is transduced over the A'$\alpha$ helix to the helix-turn-helix motif. Signal transduction is direct and can be split into 5 distinct steps. The first step is the disturbance of the FMN-binding pocket described above. These changes, specifically the opening of the $\alpha$-helices around the FMN-phospho tail, induce a movement of the B$\beta$ sheet. This can be seen in the figure 3.2.3 panel a).

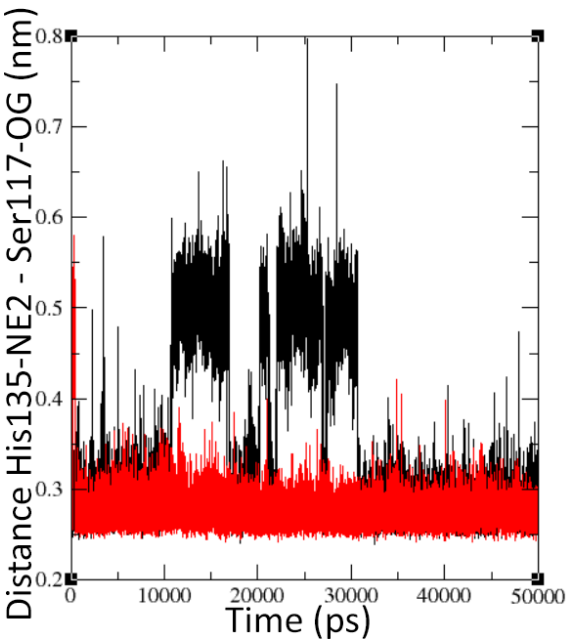


**Figure 3.2.2** The J$\alpha$-I$\beta$ anchor through the His135 and Ser117 link. Distances of the dark-dark simulation are shown in black, while the light-dark distances are represented with red color.

Panel b) of figure 3.2.3 shows the release of the A'$\alpha$ helix through the breakage of the hydrogen bond between the B$\beta$ and A'$\alpha$ regions in the light state. Figure 3.2.3 panel c) shows the signal propagation from the A'$\alpha$ helix to the HTH-motif, specifically to the J$\alpha$-helix. Finally, in figure 3.2.3 panel d) the subsequent step of the signal propagation within the HTH-region is shown. This final step is the closing of the J$\alpha$ and K$\alpha$ helices and the change in the HTH-motif geometry.

Changes in the geometry of the HTH-motif are shown in figure 3.2.4. Panel a) shows the changes in the angles between the axes of the two helices. Axis of a helix was determined by averaging the direction of backbone atoms $C_i \rightarrow O_i \rightarrow N_{i+4}$. Meanwhile, panel b) shows a simplified representation of the induced changes and the closing of the two helices, J$\alpha$ and K$\alpha$, that form the HTH-motif, upon blue light activation.

Figure 3.2.2 shows the binding of the J$\alpha$ helix to the I$\beta$ sheet through the link between the His135 and Ser117 residues. The J$\alpha$-I$\beta$ anchor persists in both the dark and light states, leaving the J$\alpha$ helix bound to the core even after the activation of the LOV domain.

The analysis of the distance between FMN-C4a and Cys55-S is shown in figure 3.2.5. Distances for each of the time points were extracted from the dark-dark simulation and grouped into distance bins. Resulting histogram shows a bimodal distribution. Integrating the surface bellow each peak of the histogram results in the ratio of the surfaces of 63% to 37%. Panel a) shows the trajectory of the inter-atomic distances between the Cys55-S and FMN-C4a atoms, while panel b) shows the histogram analysis of the trajectory.

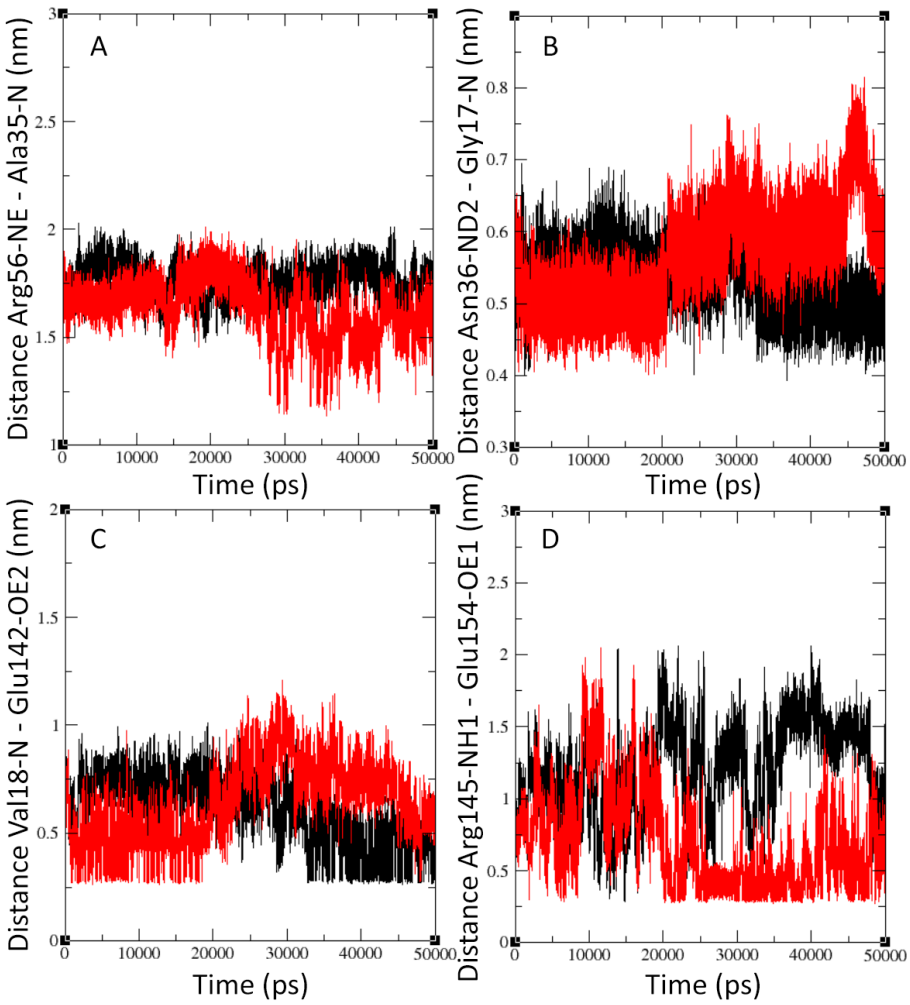


**Figure 3.2.3** Distinct signal transduction and light-activation steps. Characteristic inter-atomic distances between (A) Arg56 and Ala35, (B) Asn36 and Gly17, (C) Val18 and Glu142 and (D) Arg145 and Glu154 residues. Distances of the dark-dark simulation are shown in black in all graphs, while the light-dark distances are represented with red color.

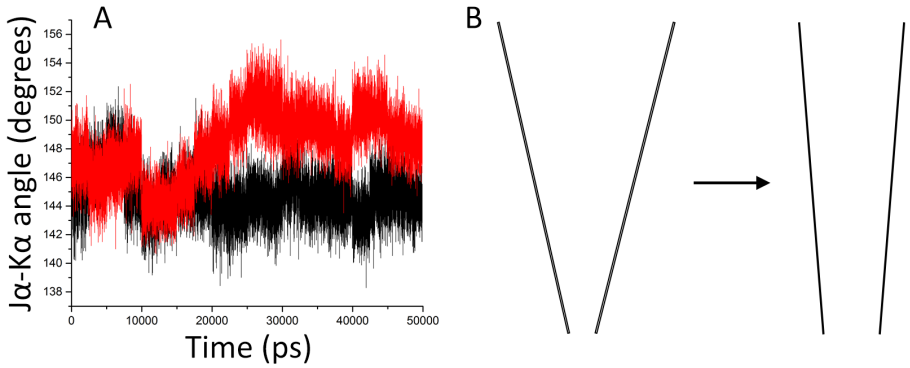


**Figure 3.2.4** Geometry changes of the HTH-motif. (A) Changes in the angles between the J$\alpha$ and K$\alpha$ axes. (B) Simplified representation of the HTH-motif changes upon activation.

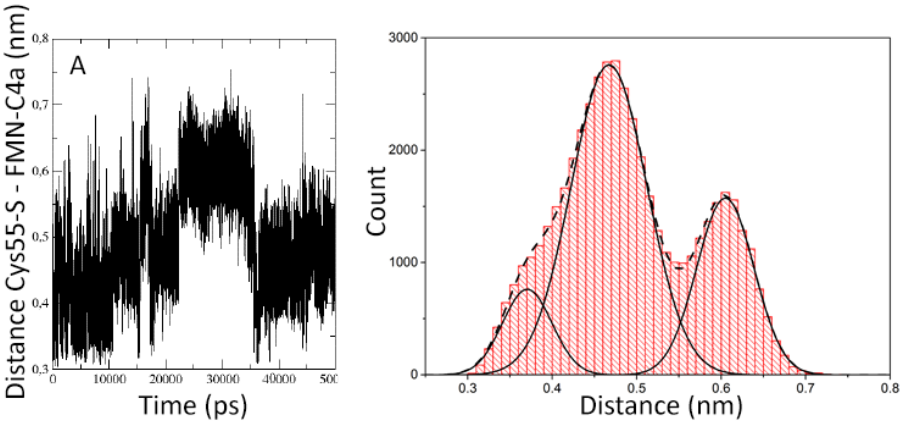


**Figure 3.2.5** Distribution of the Cys55 and FMN distances (A) Inter-atomic distances between the Cys55-S and FMN-C4a atoms of unit A in the dark-dark dimer over the simulation time. (B) Comparison of distribution functions of the distances between Cys55-S and FMN-C4a extracted from each time-point of the dark-dark trajectory.

Changes in electrostatics, upon RsLOV activation, of units A and B are shown in figure 3.2.6, where the dark state units A and B are shown in panels a) and b) and light state units A and B are shown in panels c) and d), respectively. Most noticeable is the region of the J$\alpha$ and K$\alpha$ helices, which is positively charged in unit A. The change upon blue-light activation is the increase of the positively charged region of the HTH-motif in the light state, when compared to the dark state. The J$\alpha$ and K$\alpha$ helices of the unit B are oppositely charged to the HTH-motif surface of unit A. These oppositely charged regions further strengthen the interaction between the two subunits, through electrostatic contacts. Furthermore, these electrostatic contacts allow for signal propagation beyond the single domain of RsLOV upon activation.

### 3.2.2 Dimer constructs: light-light and symmetrical dark-dark dimers

Structure of the symmetrical dark-dark construct has been generated by an in-place replacement of the unit B with a copy of the unit A. The initial structures for light-light simulation were generated by inducing the adduct state in both subunits, by forming the covalent bond between FMN-C4a and Cys55-S atoms in both units.

Figure 3.2.7 compares the differences in behavior between the light-light and symmetrical dark-dark dimers. Specifically we compare the signal propagation and mechanism steps differences. In panel a) the difference between dark and light behavior of the Arg56 and Gln49 residues is shown. Panel b) shows the distances between Arg56 and Gly48, while panel c) shows the characteristic Arg56 and Ala35 inter-atomic distances. The distances between J$\alpha$ and K$\alpha$ are shown in panel d).

In figure 3.2.8 we compare the center of mass distances between the two units forming the dark-dark, light-dark and light-light dimers over the simulation time. These

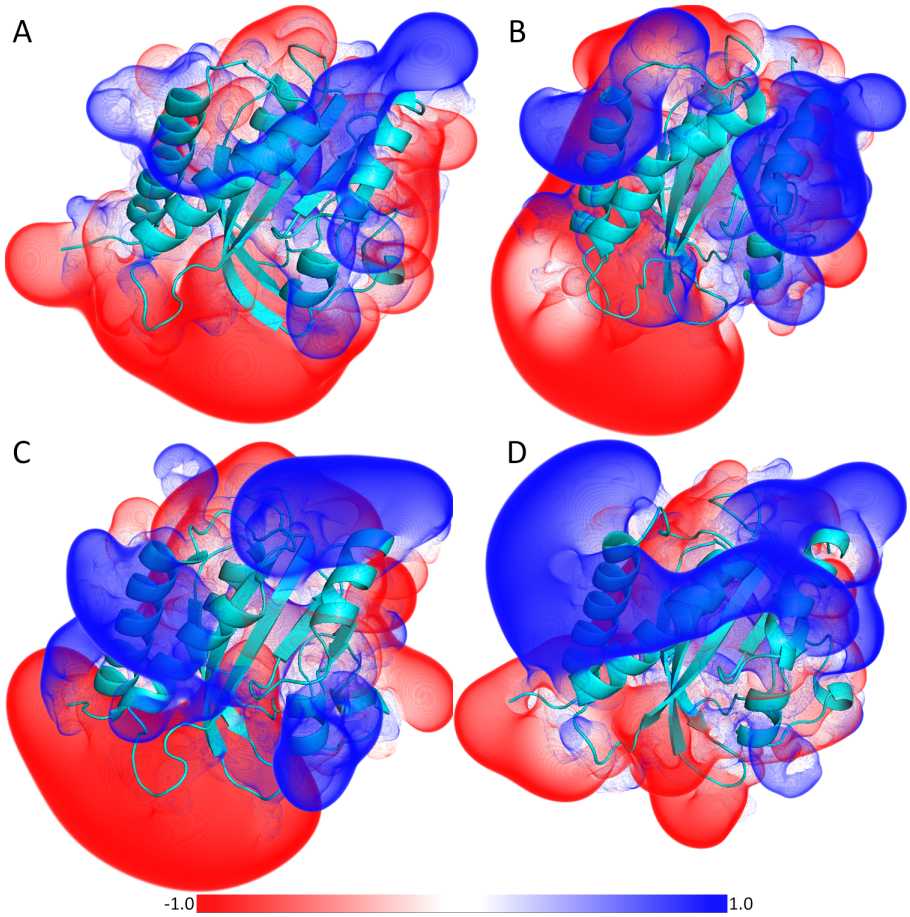


**Figure 3.2.6** APBS surfaces of RsLOV 50 ns structures of (A) dark unit A, (B) dark unit B, (C) light unit A and (D) light unit B.

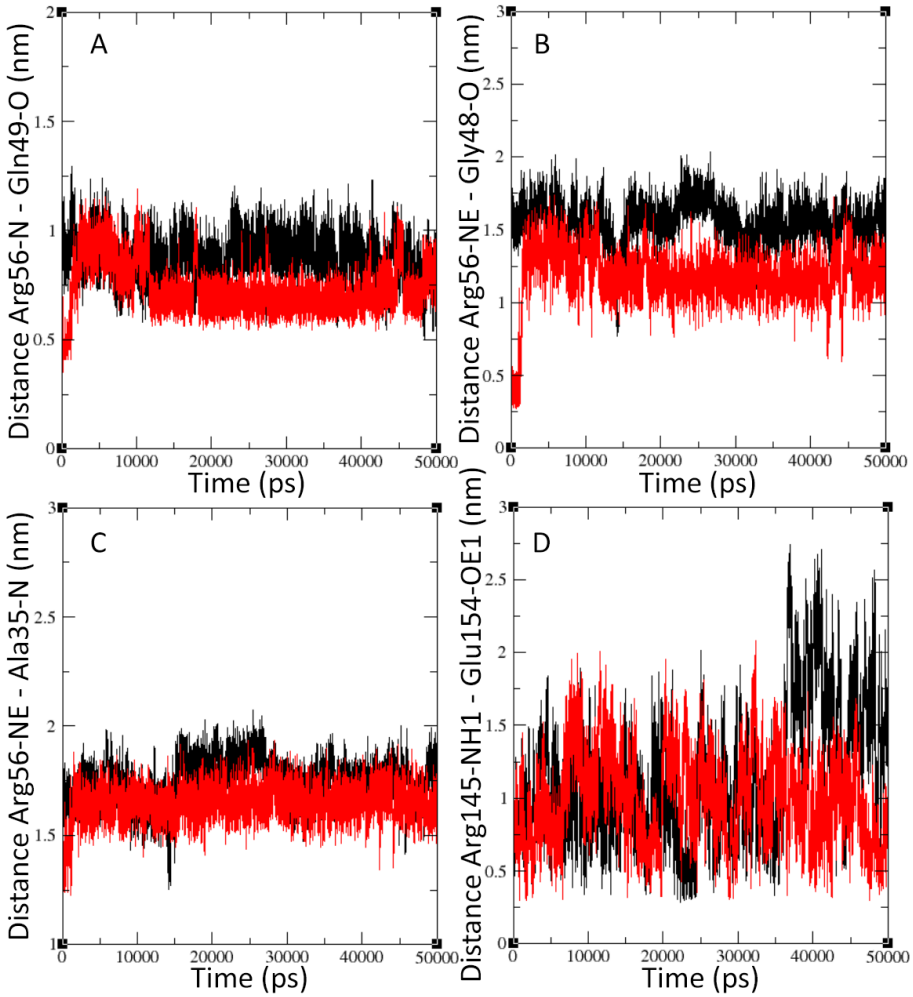


**Figure 3.2.7** Signal propagation in unit A of light-light and symmetrical dark-dark dimer simulations. Characteristic inter-atomic distances between (A) Arg56 and Gln49, (B) Arg56 and Gly48, (C) Arg56 and Ala35 and (D) Arg145 and Glu154 residues. Distances of the symmetrical dark-dark dimer simulation are shown in black in all graphs, while the light-light dimer distances are represented with red color.

distances show the differences in signal propagation past the initially-activated unit between the dimers. They also show the differences in unit-separation behavior upon blue-light activation.

To further analyze the states of different constructs we generated a density plot of radius of gyration of the HTH-motif (J$\alpha$-K$\alpha$ region) versus the Arg145 and Glu154 distance and binned the values with a 0.05 Å increment. The color scales are proportional to the density of the points for a particular bin.

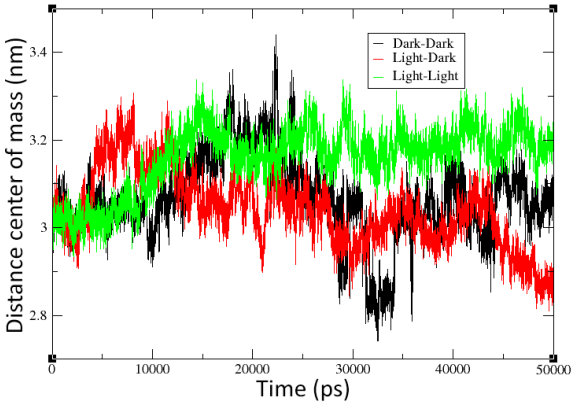


**Figure 3.2.8** Distances of centers of mass of dark-dark, light-dark and light-light dimers simulations. Dark-dark dimer distances are shown in black, light-dark distance in red and light-light distances are colored green.

### 3.2.3 Monomers

Initial structures for monomer simulations were generated by taking only the unit A of the wild type dimer RsLOV structure. For the light monomer a covalent bond was formed between FMN and Cys.

Figure 3.2.10 shows the differences in behavior between the dark and the light monomers. Panel a) shows the changes in the FMN-binding pocket through the distance between FMN and the Gln59 residue. In panel b) the distance between Arg56 and Gln49 is shown. Panel c) shows the Val18 and Glu142 distances while panel d) shows the changes induced in the HTH-motif.

Next to the the dark and light monomer simulations, we performed a simulation of the dark monomer after the light-dark simulation has completed. The behavior of this monomer is shown in figures 3.2.11 and 3.2.12. Figure 3.2.11 shows the behavior of the FMN-binding pocket in panels a)-c) while panel d) shows the J$\alpha$-I$\beta$ anchoring point through the interactions between the His135 and Ser117 residues.

Furthermore, the distances of the signal propagation steps are shown in figure 3.2.12. Panel a) shows the distance between Arg56 and Ala35, panel b) the distance between Asn36 and Gly17 while panel c) shows the distances between Val18 and Glu142. Finally, panel d) shows the changes in the J$\alpha$ and K$\alpha$ region.

Additionally in figure 3.2.13 we compare the different conformational states occupied by the two subunits of RsLOV taken after the simulations have completed. The anchoring point between the J$\alpha$ helix and the I$\beta$ sheet is emphasized with the representations of His135 and Ser117.

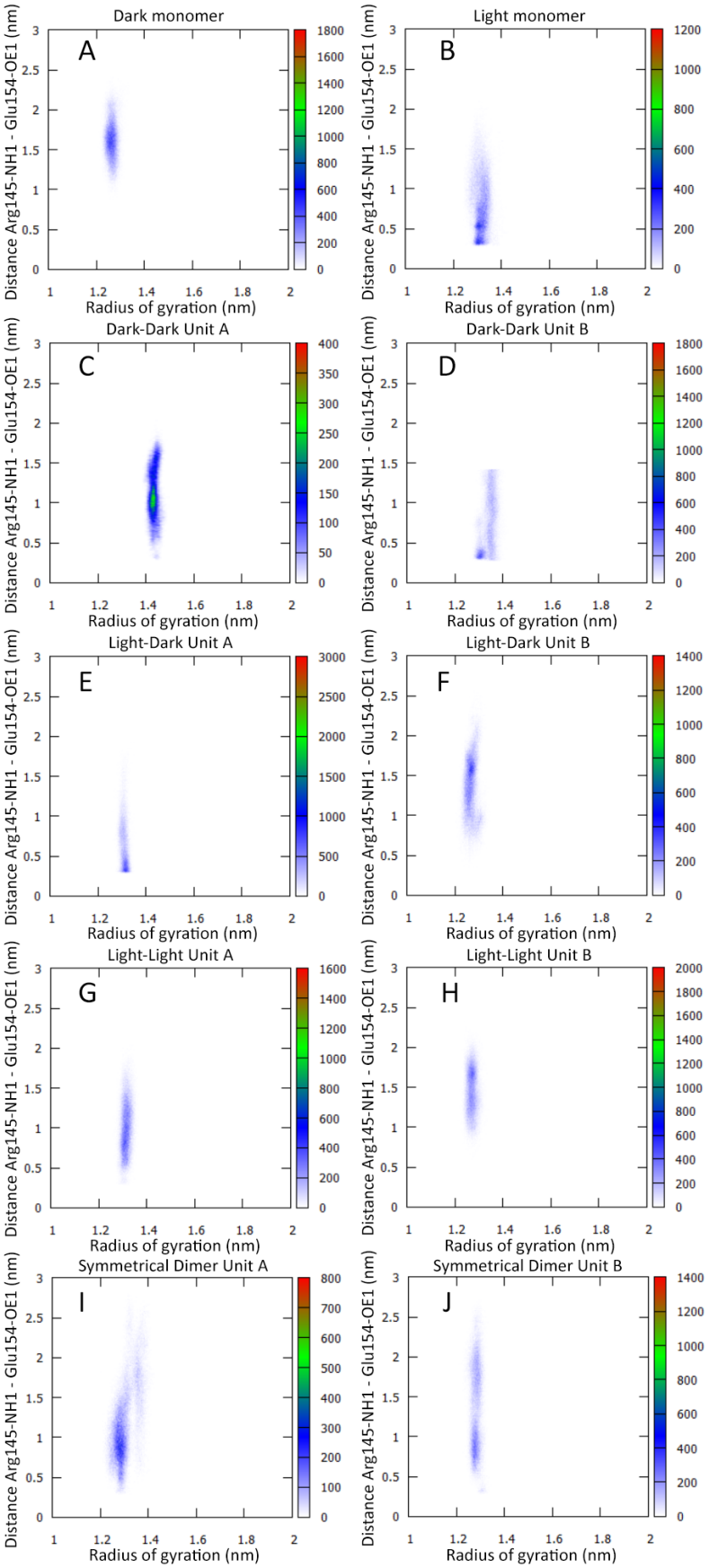


**Figure 3.2.9** Distance between the Arg145 and Glu154 residues from LOV domain as a function of the radius of gyration of J$\alpha$ and K$\alpha$ regions (HTH-motif) in the dark and light-states of different constructs. The color scales are proportional to the number of counts for the specific state observed during the simulation run.

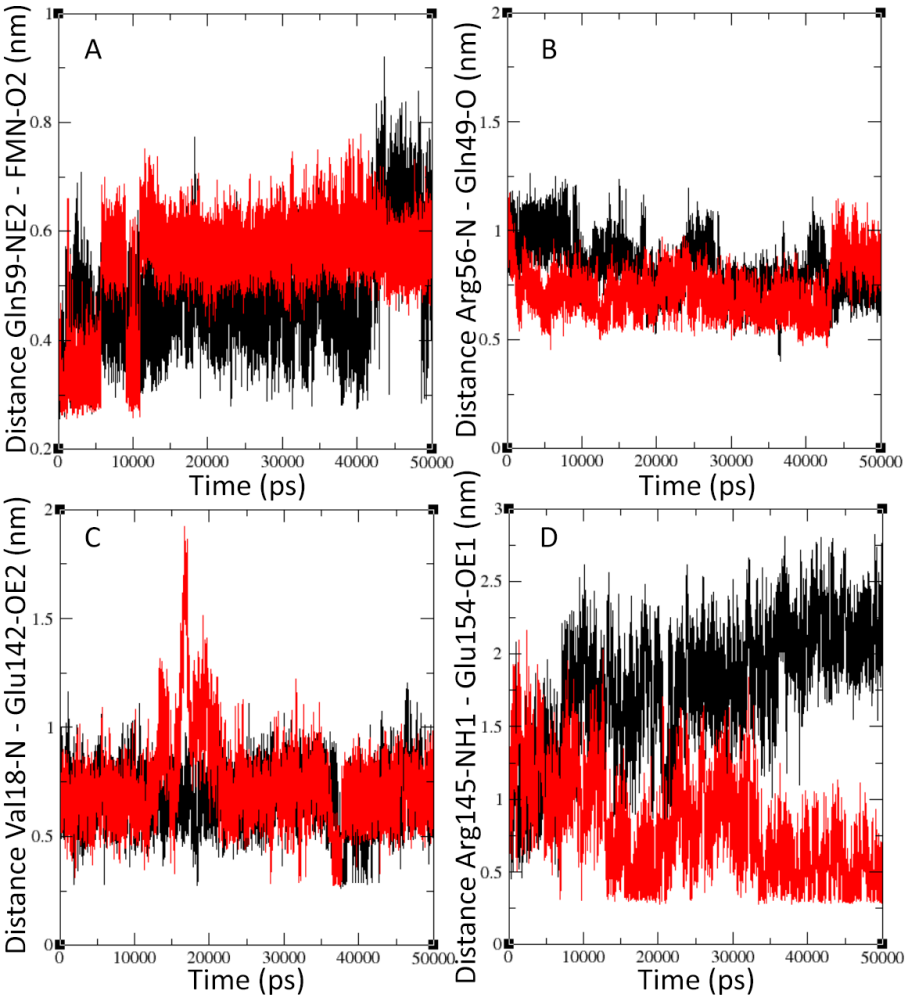


**Figure 3.2.10** Stress propagation in dark and light monomer simulations. Characteristic inter-atomic distances between (A) FMN and Gln59, (B) Arg56 and Gln49, (C) Val18 and Glu142 and (D) Arg145 and Glu154. Distances of the dark monomer simulation are shown in black in all graphs, while the light monomer distances are represented with red color.

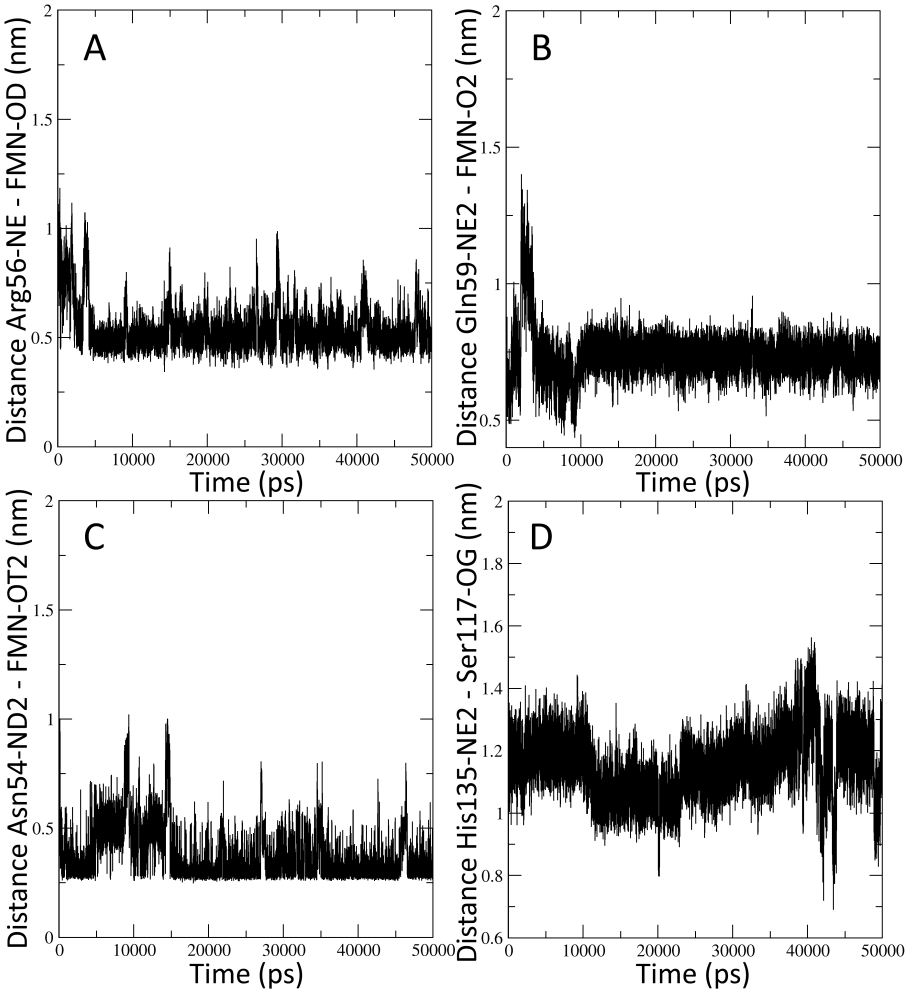


**Figure 3.2.11** FMN-binding pocket disturbance upon light-activation in the second, dark unit after the light-dark simulation has completed. Characteristic inter-atomic distances between (A) FMN and Arg56, (B) FMN and Gln59, (C) FMN and Asn54 and (D) His135 and Ser117.

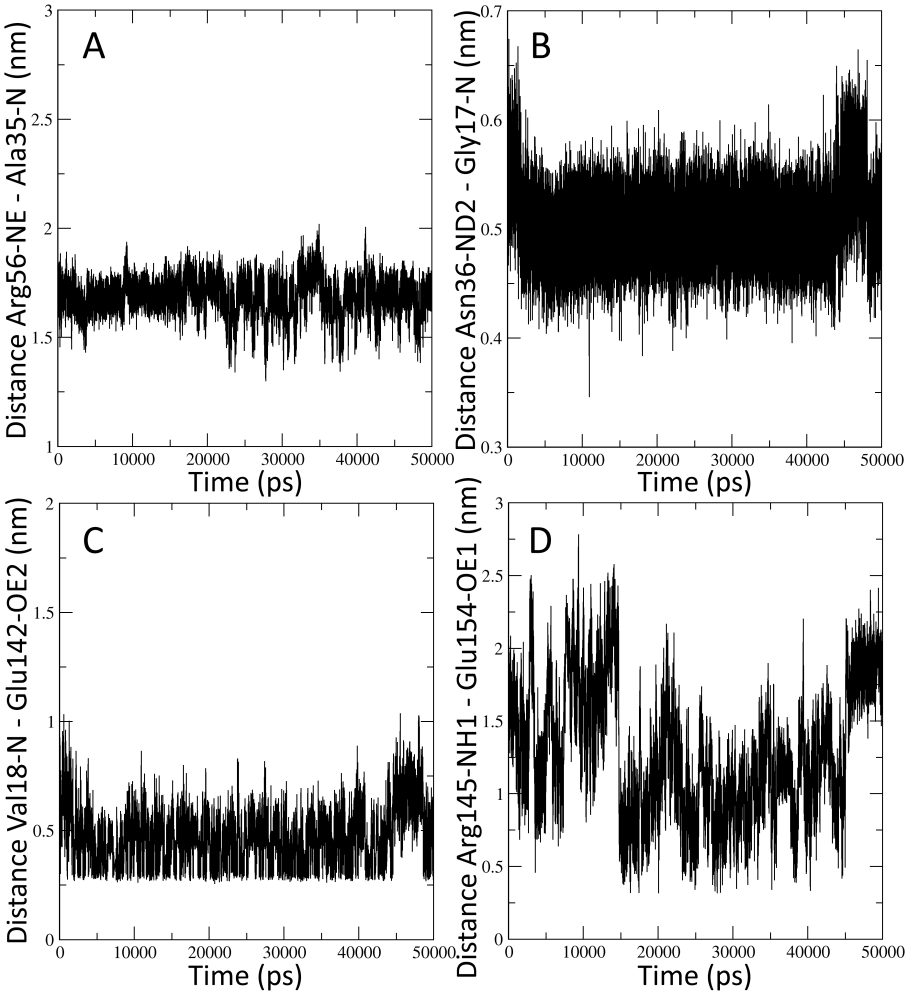


**Figure 3.2.12** Stress propagation in the second, dark unit after the light-dark simulation has completed. Characteristic inter-atomic distances between (A) Arg56 and Ala35, (B) Asn36 and Gly17, (C) Val18 and Glu142 and (D) Arg145 and Glu154.

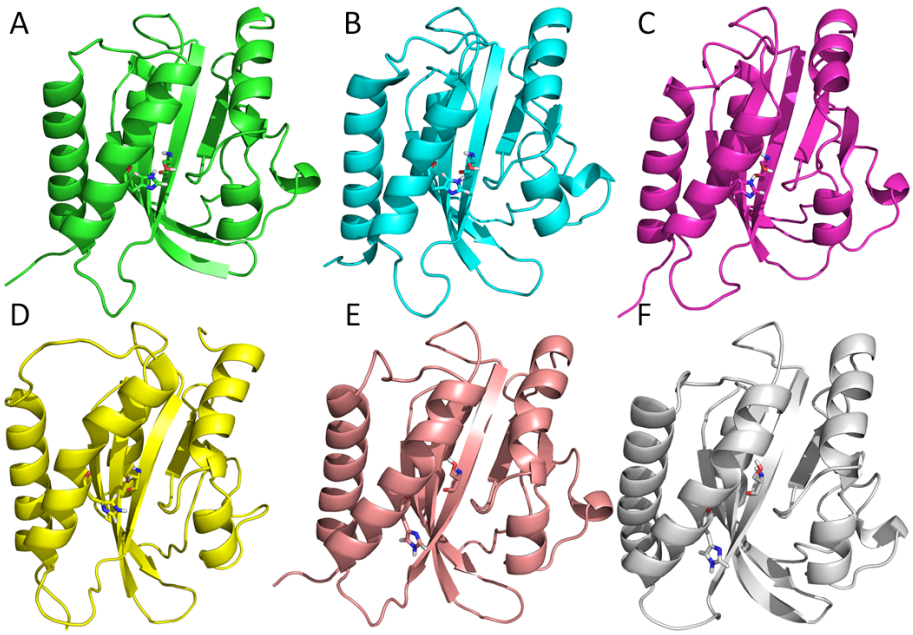


**Figure 3.2.13** Comparison of distinct structural states occupied by different units of different 50 ns simulations. (A) Chain A from dark-dark simulation, (B) chain A of light-dark simulation, (C) chain A of light-light simulation, (D) Chain B from dark-dark simulation, (E) chain B of light-dark simulation and (F) chain B of light-light simulation.

## 3.3 Discussion

Formation of the FMN-Cysteinyl adduct is a conserved first step in the light-triggered activation of LOV domains. This response to blue-light illumination causes perturbations in the flavin-binding pocket with significant changes in the residues surrounding the FMN cofactor. This is then followed by signal propagation over the A'$\alpha$ helix over to the J$\alpha$ helix, and with it to the effector domain or protein partner.

In their study, Conrad and coworkers [52] observed the flavin-cysteinyl adduct in only one of the two units and even in that one unit the light-state was formed approximately 60% of the time. In the other 40% of the time it occupied an alternative conformation. Considering that the adduct state is formed from the dark-state, we extracted the distances between the FMN-C4a and Cys55-S atoms, the two atoms forming the covalent bond in the adduct state, from the dark-dark simulation. We grouped them into bins of 0.01 nm width. Resulting histogram shows a bimodal distribution. The distances can then be grouped into two distinct groups. The data were fit using a Gaussian distribution function. Integrating the surface bellow each of the peaks in the histogram results in the surface ratio of 63% and 37% for the two groups. This points to the same behavior of the dark-dark dimer as shown in the experiments [52]. These results indicate that the dark-dark dimer spends 63% of the simulation time in conformations that are able to form an adduct. In the other 37% of the time however, the dimer occupies alternate dark-state conformations that are

unable to form an adduct as the distances between FMN-C4a and Cys55-SG are unfavorable. This is visualized in figure 3.2.5.

In RsLOV, the changes induced by the adduct formation cause an opening of the $\alpha$-helices around the phospho-tail of FMN as the first consequence of the stress induced by the adduct formation. In figure 3.2.1, the opening of the D$\alpha$ and E$\alpha$ helices is shown. Opening of the region of the E$\alpha$-F$\alpha$ helices and the FMN cofactor causes closing of the loop linking the D$\alpha$ and E$\alpha$ helices together, and formation of a link between these two regions. Additionally, the F$\alpha$ helix also shifts away from thephospho-tail of FMN, leaving the FMN-tail more exposed to the solvent.

An additional crucial step in LOV domain-signaling is the flip of the flavine-adjacent glutamine residue, conserved in most LOV domains. This flip has been previously described in the literature [16, 17, 61, 65, 66] as a key step in signal propagation. We see the Gln-flip in our light state simulations, where the changes in hydrogen-bonding network propagate through the LOV domain.

These structural changes in the binding pocket cause a movement of the anti-parallel $\beta$ sheet moiety, most notably the B$\beta$ sheet, as seen from figure 3.2.3a. The link between the B$\beta$ sheet and A'$\alpha$ helix is broken in the light state simulations, initiating the release of the A'$\alpha$ helix from the two $\beta$-sheets. This is followed by signal propagation to the J$\alpha$ helix and with it to the HTH-motif. Initially, there is a hydrogen bond between the two A'$\alpha$ and J$\alpha$ flanking helices, as shown in figure 3.2.3c. The final step of the signaling pathway within the RsLOV domain, shown in figure 3.2.3d, is the closing of the HTH-motif. Our simulations show the same closing of the J$\alpha$ and K$\alpha$ helices as reported from previous experimental work. [52] This closing of the J$\alpha$ and K$\alpha$ helices is initiated by the formation of the salt bridge between the $\text{Arg145}_{J\alpha}$ and the $\text{Glu154}_{K\alpha}$ residues. Contrary to the Conrad *et al.* [52] analysis of the crystal structure, we see no salt bridges formed between Arg145 and Glu142.

To further support the mechanism seen for the RsLOV wildtype (WT), we performed additional dark and light simulations of both dimers and monomers of a number of RsLOV mutants. Specifically, we performed simulations of RsLOV T21V, L32V, A136Y, A138Y, and R145C mutants. Mechanism events described for the RsLOV WT are consistently seen in these mutant simulations. Figure 3.3.1 schematically shows the mechanism events seen throughout RsLOV WT and mutant simulations.

The J$\alpha$ helix partially looses its helical structure but it also shows breakage and unwinding of its geometry. This is an important step in signal transduction, as J$\alpha$ detachment and unfolding has been previously [16, 67–70] described as a key component in various LOV-systems. The difference in behavior of the light-state monomer, compared to the dimer is most obvious here, when analyzing the behavior of the J$\alpha$ helix. As in the light-monomer case the breakage and unfolding of the J$\alpha$ helix is more prominent when compared to the activated dimers. The reason behind this effect is the sterical hindrance exhibited by the HTH-region of the second subunit in the RsLOV dimers.

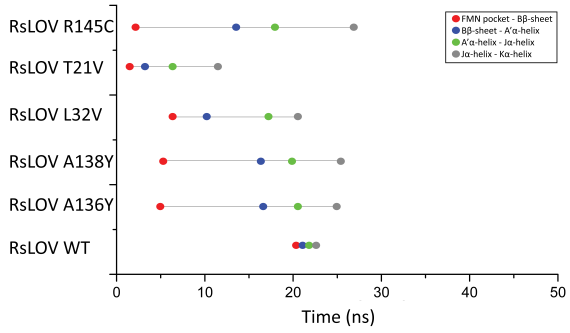


**Figure 3.3.1** Schematic representation of mechanism events seen in RsLOV WT and mutant simulations.

The salt bridge formation between the two residues, Arg145 and Glu154 of J$\alpha$ and K$\alpha$ helices respectively, causes a large geometry change in the two helices. The geometry of the HTH-motif changes from V-shaped to a more parallel-shaped geometry. This change in geometry of the HTH-region can play a role in the selection and control of the binding to the protein partners, after the dimer dissociation of the RsLOV. As the V-shaped HTH-motif can bind to protein partners and form a dimerization surface with another RsLOV unit, a more parallel J$\alpha$-K$\alpha$ region can bind to DNA or other protein partners regulating signaling pathways. Furthermore, as the HTH-regions of the two subunits form a dimerization surface for RsLOV, the changes in its geometry also play a role in the dissociation of the two RsLOV units as well as in signal transduction transfered beyond the initially-activated unit. Moreover, the changes in electrostatic potentials upon activation of RsLOV further associate the signal transduction pass beyond the first-activated unit. The APBS surfaces, shown in figure 3.2.6, show the changes in electrostatic potentials of the J$\alpha$ and K$\alpha$ helices of both units in both the dark and light states. In unit A this HTH-region is positively charged. Upon blue-light activation there is an increase of the positively charged area in the HTH-motif. The oppositely charged HTH-motif of unit B, compared to unit A, further strengthen the interactions between the two domains, and the changes in electrostatics of one HTH-region induce changes in the second HTH-motif, which ultimately allow for the signal propagation beyond the edges of a single unit. Described changes of RsLOV are shown in figure 3.3.2.

These steps of the signal transduction pathway are further corroborated when looking at the behaviors of light-light and symmetrical dark-dark dimers as well as the dark and light simulations of mutant dimers and monomers. Starting from the FMN-binding pocket perturbations, there are clear differences in behavior between the two constructs. As already described in the comparison between dark-dark and light-dark dimers behaviors, the perturbation around flavin induces a shift in the anti-parallel $\beta$-sheet region. From there the signal is transfered to the A'$\alpha$ helix and subsequently to the J$\alpha$ helix. With the J$\alpha$ helix receiving the signal, the K$\alpha$ helix is also affected, thus changing the geometry of the HTH-motif in the same fashion as with the previously described dimers.

When comparing the distances of centers of masses of the two units, between the dark-dark, light-dark and light-light dimers there are clear trends that can be identified. While, on average, the distance between the two units in the dark-dark dimers stays constant over time, surprisingly the units of the light-dark dimer come closer together over time. In contrast to the behavior of light-dark dimer subunits, the two units of the light-light dimer separate over time. The center of mass distances point towards the separation of units occurring only after the activation of both subunits in RsLOV dimers. This is not the case in light-dark dimers and the distances point to a tighter binding and an association of the units. However, because the simulation times are too short for the completion of these processes, we can not see the long-time association properties of the activated dimers in our simulations.

Conformational analysis, shown in figure 3.2.9, shows the most prevalent conformations occupied by the protein in different states and different constructs. Here we analyzed the distance of the aforementioned salt bridge between the J$\alpha$ and K$\alpha$ helices, plotted against the radius of gyration for the HTH-motif. When the unit A is in dark state, it is mostly found in a single conformation. However, when this unit is in the active state it shows several conformational clusters. Interestingly, the unit B shows a behavior that is more similar to that of the activated RsLOV unit and not to that of the dark-state unit. Unit B, in both dark and light states, shows multiple clusters of conformational states. This points towards the existence of a pseudo-lit state in RsLOV. Specifically, unit B, exhibits this behavior which is in-between the dark and the light states of the RsLOV units. The existence of these pseudo-lit states for specific mutants of certain LOV domains has been documented before. [51, 71–75] However, this has not been shown before for RsLOV wild type.

The distinct behaviors of the differently-activated dimers, taken together with the conformational analysis, lead us to the hypothesis of the varying behavior of the second unit based on its activation state, and the influence it exerts on the entire RsLOV dimer. Interestingly, there is a mutual stabilization effect between the two subunits. Unit B exhibits an influence on unit A, driving it to occupy specific conformations. In reverse, upon activation of the unit A and signal propagation to the dimerization surface and the second unit, unit B takes a largely different conformational state governed by the current state of unit A. These distinct conformations of the two units are shown in figure 3.2.13, with the anchor point between the J$\alpha$ helix and the I$\beta$ sheet emphasized. This link between His135 and Ser117 is distinct in unit B, compared to unit A, when analyzing all simulations. In the unit A this anchor between the two regions is stable over time and it does anchor J$\alpha$ helix to the RsLOV-core. This anchor persists in both dark and light states, leaving the J$\alpha$ helix bound to the protein core even after the activation of RsLOV. In unit B however, this is not the case. The hydrogen bond between Ser117 and His135 is broken, allowing the release of the J$\alpha$ helix from the rest of the protein.

Analysis of the monomer simulations leads to similar conclusions. When comparing dark and light monomers, there is a clear difference in the behavior of the two monomers. Starting from the FMN-pocket, the perturbations upon adduct for-

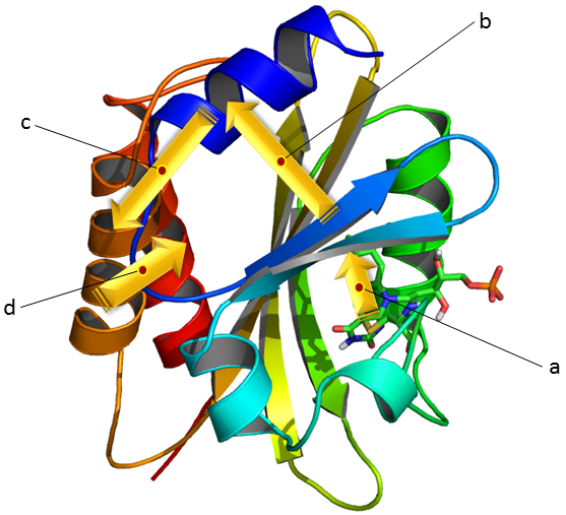


**Figure 3.3.2** The RsLOV structure after 50 ns simulation. RsLOV is shown with cartoon representation while FMN is shown with sticks. Induced structural changes are shown with yellow arrows pointing in the direction of the change going from the point of origin, the adduct formation, to the HTH-motif. Signal propagation a) from FMN pocket to B$\beta$ sheet, b) from B$\beta$ sheet to A'$\alpha$ helix, c) from A'$\alpha$ helix to J$\alpha$ helix, and d) within the HTH motif.

mation are present and are directed in the same manner as in the dimer cases. Furthermore, the signaling pathway is the same, although exhibited much faster compared to the activated dimers. This can be explained by the absence of the second unit in these structures. This absence leads to the less sterically hindered unit of the dark and light monomers, allowing for uninhibited relaxation and faster signal transduction.

Additionally, we performed the simulation of the unit B as a monomer, taken from the 50 ns snapshot of the completed light-dark simulation. The reason behind this simulation was to further support the hypothesis of the pseudo-lit state of the second unit in RsLOV dimers. Analysis of the behavior of this monomer simulation points towards geometrical changes occurring in this dark monomer during the simulation. The geometry changes induced govern the monomer into the conformational state already seen in the dark-monomer simulation. This 'return' of the second unit into the structural state of the first unit, upon the release from the dimer, further supports the hypothesis of two distinct dark-states in the units of RsLOV dimers.

Changes of the structures discussed here can also be seen in the secondary structure analyses of all RsLOV trajectories. These analyses are shown in appendix, figures A.1 to A.7.

## 3.4 Conclusions

The current study clearly demonstrates the early signaling mechanism of RsLOV upon blue-light activation. Furthermore, this study shows the transduction pathway of the signal beyond the activated unit in the RsLOV dimers and the changes induced on the dimerization surface upon passing of the signal. It also includes a proposal of the association properties of the activated dimers. Finally, the study also demonstrates the alternative dark-state conformation in the second subunit of the dimers and the mutual stabilization exhibited by the two subunits of RsLOV.

The proposed mechanism is fast, propagating from the perturbations of the flavin-binding pocket, over the anti-parallel $\beta$-sheets, to the A'$\alpha$ helix. As this helix looses its helical structure, it transfers this signal to the J$\alpha$ helix, through the mutual direct contacts. The dimerization surface, made of the J$\alpha$ and K$\alpha$ helices, is affected by this induced stress and this causes the closing of the helix-turn-helix motif.

The signal is further propagated to the second unit through the dimerization surface, and it induces structural changes in the second unit of the dimer. Subsequently, the second unit occupies the alternative dark-state conformation.

The signal induces significant changes in the dimerization surface. We see evidence for a mutual stabilization of the two units being exerted through these changes of the helices forming the interaction surface.

# 4 Switching from adduct formation to electron transfer in a light-oxygen-voltage domain containing the reactive cysteine

This chapter has been published in its entirety in *Phys. Chem. Chem. Phys.* **2017** and it is presented here as a whole. All experimental work described here has been performed by Kathrin Magerl. Supplementary information to this article is presented in the Appendix.

## 4.1 Introduction

Flavin-binding photoreceptor proteins are ubiquitous in nature. In response to blue light they regulate various biological functions in plants, bacteria, fungi and algae. [47, 76–78] They can be divided into three distinct classes: i) LOV (light-, oxygen-, voltage-sensitive) domains, ii) BLUF (blue light sensors using FAD, flavin adenine dinucleotide) domains and iii) CRY (cryptochromes).

LOV domains are the photosensory units of phototropin (Phot), harboring a non-covalently bound flavin mononucleotide (FMN) as a chromophore [79]. Upon blue-light excitation LOV domains undergo a reversible photocycle via the excited triplet state of FMN resulting in the formation of a covalent bond between the C4(a) atom of FMN and a nearby Cys [80]. This flavin-cysteinyl adduct represents the active, signaling state of the protein inducing downstream signaling within Phot. The adduct fully reverts in the dark on the timescale of seconds to hours depending on the particular LOV domain [67, 69]. Several mechanistic proposals for this adduct formation have been debated in the literature [4, 7, 10]. Of these, a radical pair mechanism seems to be most likely: the triplet state of FMN initiates hydrogen transfer, or a sequence of electron and proton transfer, from the reactive Cys resulting in the biradical pair $FMNH^{\bullet}$ $^{\bullet}SCys$ which subsequently combines to yield the adduct [81, 82]. However, although the neutral semiquinone radical $FMNH^{\bullet}$ could be detected in a LOV domain of *C. reinhardtii* [11], there is so far no unequivocal proof that it is an intermediate in the photocycle [12].

In contrast, the FAD chromophore of BLUF domains and CRY undergoes photoreduction via electron transfer upon light excitation [83–85]. The different primary activation mechanisms of these photoreceptors raise the question whether the reaction type of flavin is determined by individual amino acids or governed by the whole protein environment.

Some work has been published already to address this topic. Mutational studies on BLUF domains unraveled that replacing the electron donor Tyr by the inert amino acids Phe or Ala abolishes the photoinduced electron transfer completely, and instead formation of the excited triplet state of FAD was observed [86]. After substitution of the Tyr with Trp the electron transfer remains active, forming a radical pair with Trp instead of Tyr [87]. Additionally, it has been shown that the redox potential of FAD and Tyr as well as the $pK_a$ value of Tyr can be tuned by using fluorinated Tyr, which in turn has a large influence on the reaction kinetics of electron and proton transfer [88, 89]. A different approach was adopted by Suzuki and co-workers. By sequence alignment of two crystal structures they identified the position in a BLUF domain that corresponds to that of the Cys in LOV domains regarding the distance and orientation with respect to the flavin chromophore. They placed a Cys into this position in BLUF and inactivated the electron donor Tyr via a mutation to Phe. In this mutant the UV/Vis and IR spectra taken after photoexcitation strongly resemble those of the LOV adduct state [90].

In contrast, the formation of stable and long-lived flavin radicals was observed under aerobic and anaerobic conditions in the presence and absence of external electron donors in LOV domains when replacing the reactive Cys by Ser, Gly or Ala [91–93]. However, so far LOV domains could not be made to follow an alternative reaction type other than adduct formation as long as the reactive Cys is present, albeit there is some tolerance about the position of this Cys as shown in a mutational study on the LOV2 domain of *A. sativa* [67].

Very recently, Yee *et al.* demonstrated that the LOV domain VVD and the artificial construct YF1 are capable of transducing the light-induced signal downstream within the protein although the reactive Cys is lacking, identifying the neutral semiquinone radical as the signaling state [66]. Analogously, some BLUF domains do not require the formation of a radical intermediate to retain their biological function [94]. These findings lead to the hypothesis that the primary reaction of these photoreceptors is not decisive for the initial activation of the proteins and does not modulate the second order of events. Instead, different flavin photoreactions lead to the same decisive structural modification of the photoreceptors. In BLUF domains this is rearrangement of hydrogen bonds, in LOV domains this is likely the protonation of FMN-N5, caused either by adduct formation or by electron coupled proton transfer [66]. Being able to manipulate the photoreactivity of a flavin chromophore inside a protein core and hence being able to control the timescales of activation via switching between different mechanisms within the same domain structure could therefore contribute substantially to the design and engineering of optogenetic tools and reporter proteins in living cells [95, 96]. LOV domains have already been suc-

cessfully engineered to serve as fluorescent reporter proteins as an alternative to the GFP family [97]. Additionally, in the field of optogenetics, LOV domains were applied as photoswitches in living cells as exemplified by the photoactivatable Rac1 [98].

In the present study, we changed the photoreaction of the LOV1 domain of *C. reinhardtii* via mutation of Phe41 to Tyr. Upon blue-light excitation, the F41Y mutant does not form the characteristic flavin-cysteinyl adduct but performs an electron transfer with the introduced Tyr even though the reactive Cys57 is present. When additionally mutating this Cys to Ala or Ser the yield and protonation state of the intermediate flavin radicals can be tuned. We present spectroscopic data in combination with MD simulations to unravel the primary FMN photochemistry, identify radical species participating in the reaction, and discuss structural changes, hydrogen bond rearrangements, and solvent accessibility in the proteins. These results allow us to propose key parameters important for an electron transfer to occur in LOV domains.

## 4.2 Methods

### 4.2.1 Sample preparation

The gene encoding for the cDNA of the LOV1 domain of *Chlamydomonas reinhardtii* (amino acids 16-133), cloned into the vector His-p2x as described previously [99], was used as a template DNA for site-directed mutagenesis. Oligonucleotide primer sequences are listed in the supplementary information (Table S1)†. The resulting sequences of the mutants were verified by DNA sequencing (Seqlab Sequence Laboratories, Göttingen, Germany).

### 4.2.2 Protein expression and purification

The *Cr*LOV1 mutants were expressed in *E.coli* BL21 (DE3) and purified via Ni-NTA affinity chromatography. The eluted protein was dialyzed four times against 10 mM sodium phosphate buffer, pH 8.0, supplied with 10 mM NaCl in order to remove non-bound FMN. The proteins were concentrated using centrifugal concentrators (Amicon Ultra-15, Millipore, NMWL 10 kDa) and stored in 10 mM sodium phosphate buffer (10 mM NaCl, pH 8.0) at 4 $^{o}$C in the dark.

### 4.2.3 UV/Vis Spectroscopy

UV/Vis spectra were recorded with a Lambda 9a spectrometer (Perkin Elmer, Waltham, MA, US) at 20 $^{o}$C using a temperature controlled cuvette holder.

### 4.2.4 Determination of $S_1$ lifetimes

The fluorescence lifetimes of the F41Y mutants were determined at room temperature via time correlated single photon counting (TCSPC) using the PicoHarp 300 system (PicoQuant). The samples were excited at 378 nm with a pulsed laser diode (Pico Brite, Horiba Scientific, PB-375L $\pm$ 7 nm, $\lambda_{max} = 378$ nm, pulse width $< 100$ ps, $\omega = 2$ MHz). Emission was observed at 510 nm in a 90 $^o$ geometry relative to excitation with a photomultiplier tube. The response function of the excitation source (IRF) was detected separately at 376 nm. Data analysis was performed with the FluoFit software version 4.5 (PicoQuant) using an exponential model with reconvolution for compensating IRF effects.

### 4.2.5 Determination of fluorescence quantum yields, $\Phi_F$

. Measurements of the fluorescence quantum yields, $\Phi_F$, at ambient temperatures were performed with a C9920-02 (Hamamatsu Photonics) integrating sphere system with an error of $\pm$ 5%. The samples were excited at $\lambda_{exc} = 450$ nm.

### 4.2.6 Transient Absorption Spectroscopy and data analysis

Transient absorption spectroscopy in the $\mu$s-time range was performed with a streak camera setup (C7700, Hamamatsu Photonics). The samples were excited at 447 nm via an OPO pumped by a Nd:YAG laser (Surelite II, Continuum) with a pulse width of 8 ns. The pulse energy was adjusted to be 8-10 mJ per pulse in front of the cuvette. A fused silica flow cuvette with 2 mm optical path length for excitation and 10 mm path length for probe light was used. The samples were pumped through the cuvette during the measurement using a peristaltic pump (LKB Pump P-1, Pharmacia Biosystems) in order to exchange the sample volume between two excitation cycles. The overall sample volume per measurement was around 10 mL with an optical density at 450 nm of approximately 0.5. Each measurement was performed as a sequence of 100 excitation cycles with one cycle consisting of four individual streak images: $I_{L,P}$, $I_D$, $I_P$, $I_D$. With $I_{L,P}$ being an image with laser and probe light on, $I_D$ being the dark spectrum and $I_P$ referring to probe light on only.

The 2D TA data were analyzed by global lifetime analysis with the following model:

$$\Delta A(t,\lambda) = \left( D_0(\lambda)\delta(t) + \sum_{j=1}^{N} D_j(\lambda) exp(-\kappa_j t) \right) \otimes g_{app}(t - t_0) \qquad (4.1)$$

where $\otimes g_{app}(t\text{-}t_0)$ indicates convolution with the apparatus function approximated by a Gaussian, $\delta(t)$ is the Dirac delta function, and $N$ is the number of decay components considered. The result of this fit are the rate constants $\kappa_j = 1/\tau_j$ and the corresponding decay associated difference spectra (DADS) $D_j(\lambda)$. The spectrum

$D_0(\lambda)$ accounts for the contribution of fluorescence and scattered laser light. Since the resolution of the streak camera of ca. 2.5 pixels corresponds to a time of ca. 50 ns, these contributions are not resolved. Further details about the setup and data analysis are described in Kutta *et al.*[100].

### 4.2.7 Spectral fitting of DADS

Spectral fitting of the $D_2$ of *Cr*LOV1-F41Y, *Cr*LOV1-F41Y/C57A and *Cr*LOV1-F41Y/C57S was performed using a least-square algorithm without any constraints. A linear combination of the reference spectra $FMNH^{\bullet}$ minus FMNox, $FMN^{\bullet-}$ minus FMNox and $TyrO^{\bullet}$ was used to calculate the model spectra, and the coefficients were optimized.

The reference spectrum of $FMNH^{\bullet}$W at 460 nm) and spectral changes were monitored via UV/Vis until the spectrum stayed constant. In order to obtain the difference spectrum, the ground state spectrum, FMNox, was subtracted. This is possible since the relative amplitudes of both spectra are known from the isosbestic point at 500 nm.

The reference spectrum of $FMN^{\bullet-}$ was generated by flash photolysis of a 40 $\mu$M solution of *Cr*LOV1 C57S in the presence of 10 mM BME and oxygen in 10 mM sodium phosphate buffer (10 mM NaCl, pH 8.0). It is known that under these conditions the long-lived DADS ($\tau \gg 500\,\mu$s) is the SADS (species associated difference spectrum) of the FMN radical anion [12].

For the reference spectrum of $TyrO^{\bullet}$, FMN was dissolved in 10 mM sodium phosphate buffer, pH 8.0 supplemented with 10 mM NaCl at a concentration of 40 $\mu$M, resulting in an optical density of 0.5 at 450 nm. To this solution either L-Tyr or L-Cys was added at a concentration of 100 mM. TA was measured on a 20 $\mu$s (Tyr) or on a 10 $\mu$s (Cys) time window. The difference spectrum of the $TyrO^{\bullet}$ radical was thus obtained as the difference of the spectra from the Tyr and Cys experiments, scaled to the same ground state bleach of FMNox at 450 nm.

### 4.2.8 Modelling

For protein structure predictions MODELLER version 9.11 software package [101] was used, employing the homology modelling method (comparative modelling). The model of each mutant has been generated based on the crystal structure of native CrLOV1 (pdb code: 1N9L, determined by Fedorov *et al.* [58]), with required point mutations: F41Y, F41Y C57A, F41Y C57S and F41Y C57G.

### 4.2.9 MD Simulations

To resolve the behavior of *Cr*LOV and its mutants at a molecular level, MD trajectories were generated using the GROMACS version 4.6.5 software package with the GROMOS 96-43a1 forcefield [102]. This forcefield has been shown to reliably reproduce experimental results on various protein-solution systems [103, 104]. For the simulations full particle-mesh-Ewald (PME) electrostatics were used with a Coulomb cutoff of 0.9 nm, Van der Waals interactions were computed with a cutoff of 1.4 nm. Proteins were put in a cubic box with the distance between the edge of the box and the closest protein atom set to 1.0 nm. To neutralize the system, 2 sodium ions were placed at random positions within the cubic box. The systems were equilibrated for 5 ns using the NPT ensemble at 300 K and 1 bar, where all parts of the system were coupled to the Nosé-Hoover-thermostat and Parrinello-Rahman barostat [105]. Following the equilibration phase, a production run of 20 ns non-invasive thermostating [16] was performed in which the protein and the FMN were decoupled from the thermostat, while the solvent and ions remained coupled. The non-invasive method allows the protein to follow its natural dynamics and sample configurations which are far from the equilibrium. For the numerical integration of the equations of motion, we have used the Leap-frog integrator with a timestep of 2 fs.

## 4.3 Results

### 4.3.1 Photoreaction of *Cr*LOV1-F41Y

We substituted Phe41 isosterically with Tyr in *Cr*LOV1 in order to investigate the influence of an electron donor amino acid in the protein core on the photoreactivity of the chromophore FMN. The F41Y mutant was expressed and purified like the wild type (wt) as a soluble, green protein in solution.

The steady-state UV/Vis spectrum of *Cr*LOV1-F41Y (figure 4.3.1) exhibits the typical fine structured absorption bands with maxima at 473 nm, 445 nm, 368 nm and 352 nm, blue-shifted by 3 nm compared to the wt. The absorption band at 352 nm is a good indicator for the hydrogen bonding network and the polarity in the FMN binding pocket [106, 107]. The high spectral similarity between *Cr*LOV1-F41Y and *Cr*LOV1 wt indicates that the mutation did not lead to major structural changes within the protein. In contrast to *Cr*LOV1 wt, the steady-state UV/Vis spectrum of *Cr*LOV1-F41Y does not change upon blue light illumination, meaning no long-lived adduct is formed in this mutant. To further investigate the behavior of *Cr*LOV1-F41Y upon blue light excitation, transient absorption (TA) spectroscopy in the $\mu$s time range was applied. Figure 4.3.2 shows difference absorption spectra of *Cr*LOV1-F41Y at

200 ns and 4.9 $\mu$s after laser excitation. For comparison, previously published spectra of *Cr*LOV1 wt are included [12]. All spectra are scaled to the same minimum at

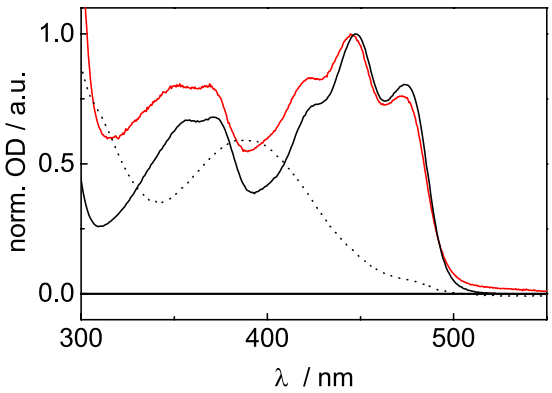


**Figure 4.3.1** Comparison of the UV/Vis spectra of *Cr*LOV1 wt in the dark (black line) and in the adduct state (dotted line) and of *Cr*LOV1-F41Y (red line). The steady-state UV/Vis spectrum of *Cr*LOV1-F41Y does not change upon blue light illumination.

447 nm. The negative absorption band at around 450 nm corresponds to the ground state bleach of the oxidized FMN, FMNox. Positive absorption signals are detected for *Cr*LOV1-F41Y from 493-745 nm and below 415 nm after 200 ns. The wt exhibits a similar spectrum blue-shifted by 5 nm. After 4.9 $\mu$s, a positive band can be detected at 374-407 nm and above 493 nm up to 660 nm in the case of *Cr*LOV1-F41Y. The positive absorption band above 660 nm almost decayed to zero at this time delay. The difference spectrum of *Cr*LOV1 wt in figure 4.3.2 after 4.9 $\mu$s is characterized by the positive absorption of the adduct state at 376-409 nm and the ground state bleach. Above 500 nm, no absorption signals contribute to the spectrum.

We performed global lifetime analysis of the TA data, two lifetime components were sufficient for a good fit. Figure 4.3.3 shows the two decay associated difference spectra (DADS) corresponding to decay time constants of $\tau_1 = 1.3\,\mu$s and $\tau_2 > 10\,\mu$s. The first DADS ($D_1$) exhibits positive signals above 490 nm and below 420 nm next to the negative ground state bleach of FMNox. The broad absorption band with the typical structure in the range 650 to 700 nm can be associated with the excited triplet state of FMN [108]. The triplet decay time of *Cr*LOV1-F41Y is almost the same as in *Cr*LOV1 wt, which was reported to be 1.16 $\mu$s [12]. The second DADS ($D_2$) is non-decaying within the observed streak window of 10 $\mu$s. This $D_2$ has positive absorption in the range of 375 nm to 403 nm and from 495 nm to 670 nm, resembling the spectrum of the neutral semiquinone radical $FMNH^{\bullet}$ [108, 109]. Measurements on a longer timescale reveal a decay time of this species of ca. 300 $\mu$s in air-saturated solution. In degassed solution the spectrum of $FMNH^{\bullet}$ was observed with constant amplitude for 15 minutes, but disappeared when air was admitted. The spectrum of this radical in the LOV environment is known from photochemical reduction experiments of *Cr*LOV1-C57G with EDTA or $\beta$-mercaptoethanol [91–93, 110]. Comparison of these spectra with $D_2$ of *Cr*LOV1-F41Y revealed a significant difference in the range 500-530 nm, indicating an additional species contributing to $D_2$. $FMNH^{\bullet}$ is formed by electron transfer to FMN, leading to the radical anion $FMN^{\bullet-}$ which is subsequently protonated. In the photoreduction experiment with *Cr*LOV1-C57G

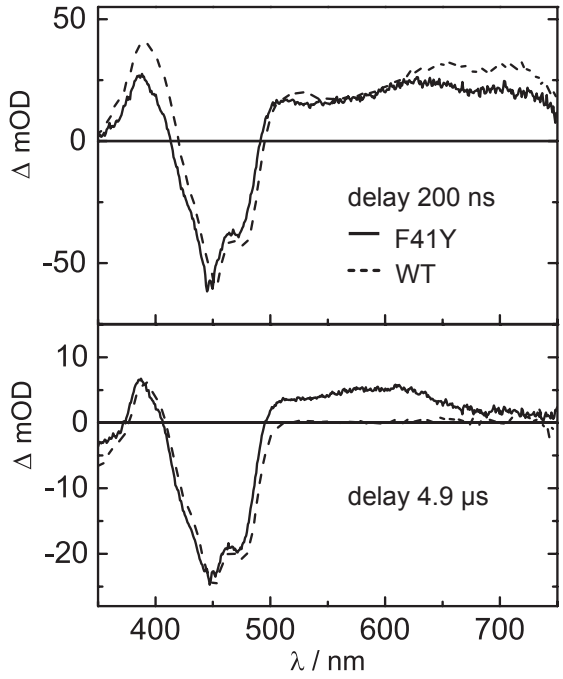


**Figure 4.3.2** Difference absorption spectra extracted from a TA measurement recorded with a 10 $\mu$s time window of *Cr*LOV1-F41Y (solid line) 200 ns and 4.9 $\mu$s after laser excitation. Spectra of *Cr*LOV1 wt (dotted line) are shown for comparison. The noise level of the data was reduced by singular value decomposition (SVD) and all spectra were scaled to the same minimum at 447 nm.

the electron was donated by $\beta$-mercaptoethanol. The cation of this donor or its decay products do not absorb in the visible spectral range and protonation of $FMN^{\bullet-}$ is complete in this steady state experiment. In the case of *Cr*LOV1-F41Y, however, the electron is most probably donated by Tyr41. The initially produced tyrosine radical cation usually deprotonates quickly to the neutral radical. On the other hand, protonation of the radical anion $FMN^{\bullet-}$ can be a slow process on the $\mu$s timescale. Hence we considered the following species for a possible contribution to $D_2$: $FMNH^{\bullet}$, $FMN^{\bullet-}$ and $TyrO^{\bullet}$.

We then tried to fit the observed $D_2$ to a linear combination of reference spectra of FMNox, $FMNH^{\bullet}$, $FMN^{\bullet-}$ and $TyrO^{\bullet}$ (for description how these reference spectra were generated see Methods part). A fit to the reference spectrum of $FMNH^{\bullet}$ minus FMNox did not lead to reasonable agreement, nor did a fit including also the reference spectrum of $FMN^{\bullet-}$ minus FMNox. However, when the reference spectrum of $TyrO^{\bullet}$ was included, a fit with good agreement over the whole spectral range of 350-750 nm could be obtained. Figure 4.3.4 shows $D_2$ (black curve in the main plot), the reference spectra of $FMNH^{\bullet}$ minus FMNox (green) and $FMN^{\bullet-}$ minus FMNox (blue), and the linear combination of these in the fit (purple). The inset shows the difference between $D_2$ and the linear combination of the flavin species (red) in comparison to reference spectra of $TyrO^{\bullet}$ in aqueous solution (black and blue lines). Apparently, the two flavin radical species yield a good description of $D_2$ in the range

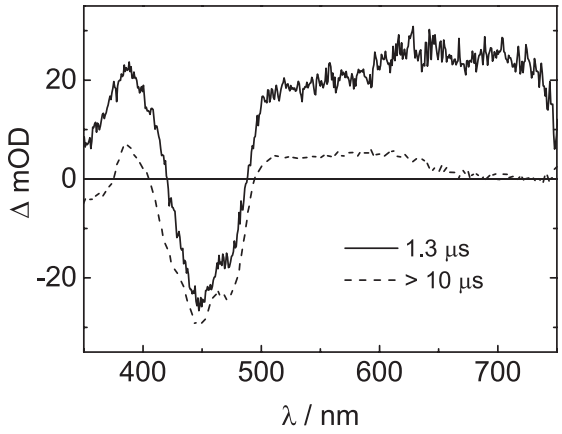


**Figure 4.3.3** DADS obtained from global lifetime analysis of *Cr*LOV1-F41Y. The first DADS (solid line) decays with a time constant of 1.3 $\mu$s, while the second DADS (dashed line) is non-decaying within the measured time window of 10 $\mu$s.

$\lambda > 500$ nm, but $TyrO^{\bullet}$ is required to match the data at 400 nm . We conclude that electron transfer in *Cr*LOV1-F41Y occurs from Tyr41 to FMN, and that the flavin radical anion is still not completely protonated after 10 $\mu$s.

## 4.3.2 Additional mutation of the reactive Cys57

No evidence for adduct formation was found in *Cr*LOV1-F41Y although the reactive Cys is still present. To elucidate, whether this Cys57 is participating in the observed electron transfer reaction of *Cr*LOV1-F41Y, we additionally replaced the reactive Cys with the polar amino acid serine, and the unpolar amino acids alanine and glycine. The UV/Vis spectra (supporting information figure S1) are in good agreement with the UV/Vis spectrum of *Cr*LOV1 wt indicating an intact hydrogen bonding network around the FMN.

We performed TA measurements at pH 8.0 on a 20 $\mu$s time window with *Cr*LOV1-F41Y/C57A and on a 50 $\mu$s time window with *Cr*LOV1-F41Y/C57S. The data were analyzed by global lifetime analysis using two exponential functions for each data set. The $D_1$ displayed in figure 4.3.5A decay with time constants of 1.7 $\mu$s for *Cr*LOV1-F41Y/C57A and 2.8 $\mu$s for *Cr*LOV1-F41Y/C57S, i.e. they are associated with the decay of the excited triplet state of FMN. $D_2$ of *Cr*LOV1-F41Y/C57S decays within 28.0 $\mu$s, while the C57A mutant is non-decaying within the measured time window (figure 4.3.5B). Both DADS in figure 4.3.5B exhibit similar spectral features of $FMNH^{\bullet}$ like *Cr*LOV1-F41Y. Interestingly, the spectral signatures of the radicals appear not only in the $D_2$, but also in the $D_1$ exemplified by subtraction of $D_1$ of the C57G mutant which is entirely due to triplet absorption (see figure 4.3.5C). The $FMNH^{\bullet}$ absorption varies significantly between the different mutants in the range of 540 nm to 655 nm. The highest amplitude can be found in *Cr*LOV1-F41Y/C57S, followed by *Cr*LOV1-F41Y/C57A and *Cr*LOV1-

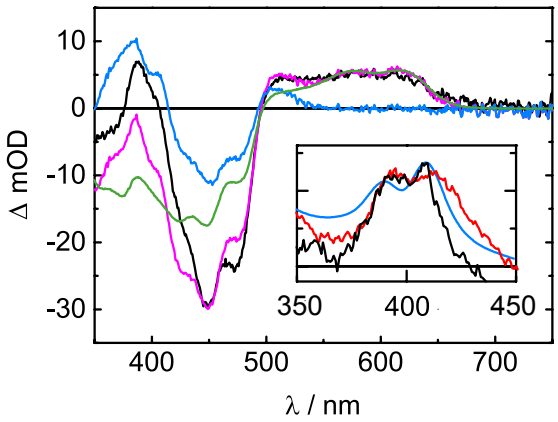


**Figure 4.3.4 Main plot:** $D_2$ of *Cr*LOV1-F41Y (black line) compared to the sum of the reference difference spectra (magenta line) for the species $FMNH^{\bullet}$ minus FMNox (green line) and $FMN^{\bullet-}$ minus FMNox (blue line). **Inset:** The red spectrum represents the difference between $D_2$ and the sum of the flavin radical species, i.e. the difference between the purple and the black line of the main plot. This spectrum resembles the spectrum of the $TyrO^{\bullet}$ radical. For comparison, reference spectra of $TyrO^{\bullet}$ in aqueous solution are included (black line: our data; blue line: reference [111]), blue-shifted by 2 nm compared to the red spectrum.

F41Y. Furthermore, the positive absorption below 405 nm differs between CrLOV1-F41Y/C57S and CrLOV1-F41Y/C57A, being more intense in the latter one with a shoulder at 398 nm. In combination with the pronounced peak around 500 nm in *Cr*LOV1-F41Y/C57A, these spectral features are characteristic for the FMN radical anion, $FMN^{\bullet-}$.

A fit of the linear combination of the aforementioned reference spectra revealed contributions of $FMNH^{\bullet}$, FMNox, $FMN^{\bullet-}$ and $TyrO^{\bullet}$ to the $D_2$ of *Cr*LOV1-F41Y/C57S (supporting information, figure S2A). In contrast, $D_2$ of *Cr*LOV1-F41Y/C57A could be best described using a linear combination of $FMNH^{\bullet}$ minus FMNox and $FMN^{\bullet-}$ (supporting information, figure S2B). Including the $TyrO^{\bullet}$ spectrum did not improve the fit. The results from spectral fitting indicate that the introduced Tyr41 is quite likely the terminal amino acid of the electron transfer chain and therefore the electron donor, since the spectrum of $TyrO^{\bullet}$ needs to be taken into account to fit the $D_2$ of *Cr*LOV1-F41Y and *Cr*LOV1-F41Y/C57S. Only in the case of *Cr*LOV1-F41Y/C57A we could not establish the contribution of $TyrO^{\bullet}$ via spectral fitting of $D_2$. But in this mutant, the $FMN^{\bullet-}$ features are more pronounced compared to the other F41Y mutants, which has a strong absorption in the same spectral region as the $TyrO^{\bullet}$ radical.

The photophysical behavior of *Cr*LOV1-F41Y/C57G is different compared to the other F41Y mutants investigated. Analysis of a 100 $\mu$s TA dataset results in two DADS (supporting information, figure S3A and B). $D_1$ exhibits the typical spectral features of the FMN triplet state and decays with $\tau = 28.0\,\mu$s, prolonged by a factor of 10 compared to the corresponding DADS of the other F41Y mutants. $D_2$ is non-decaying within the measured streak window and the spectrum exhibits high

similarity with $D_1$. This component has an amplitude of approx. 10 % of the faster decaying component. In *Cr*LOV1-F41Y/C57G no spectral evidence for FMN radicals can be observed. These findings suggest a biphasic triplet decay like already observed in the case of *Cr*LOV1-C57G [12] indicating the presence of an additional quenching channel. The same has been reported for Pp2FbFP L30M, a LOV domain based singlet oxygen sensitizer [112]. It has been suggested that two distinct triplet states are formed in those proteins which can be divergently quenched by oxygen due to varying accessibility of molecular oxygen into the binding pocket. Glycine is the least sterically demanding natural amino acid. This allows for more space in the binding pocket and therefore increased flexibility of the chromophore.

### 4.3.3 pH influence

Our TA data show that mainly the protonated $FMNH^{\bullet}$ can be observed upon blue light excitation of the *Cr*LOV1-F41Y mutants. In order to identify the proton donor in this reaction we increased the pH of the buffer solutions to more basic values up to 11.0 and repeated the TA measurements. Figure 4.3.6 compares the $D_2$ at different pH values of the respective *Cr*LOV1-F41Y mutants resulting from analysis of the TA data.

All three F41Y mutants form the neutral radical $FMNH^{\bullet}$ even at strongly basic pH, although with a reduced yield. The strongest decrease of approx. 50 % can be observed in the case of *Cr*LOV1-F41Y/C57S (figure 4.3.6B). In the $D_2$ of *Cr*LOV1-F41Y and *Cr*LOV1-F41Y/C57A (figure 4.3.6A and C) the spectral features of $FMN^{\bullet -}$ become more pronounced with increasing pH including the absorption bands around 500 nm and below 400 nm. The fact that $FMNH^{\bullet}$ can still be observed at strongly basic pH indicates that the proton donor must be protonated as well up to a pH of 10.0-11.0. Therefore, protonation by the solvent is not possible but must be performed by an amino acid of the LOV core. According to the crystal structure of *Cr*LOV1 (pdb 1N9L) as well as our MD simulations, the only proton donor in the binding pocket of the *Cr*LOV1-F41Y mutants is Tyr41. In the case of *Cr*LOV1-F41Y also Cys57 is possible. The $pK_a$ value of free Cys is 8.18 and the one of Tyr is 10.07. However, the $pK_a$ values of amino acids can be significantly shifted by the protein environment [113]. Hence, we calculated the $pK_a$ value of Tyr41 in the *Cr*LOV1-F41Y mutants using the $H^{++}$ program v3.2 [114] resulting in $pK_a$ values of 11.45 (F41Y C57S) and $>12$ (F41Y, F41Y/C57A). Cys57 in *Cr*LOV1-F41Y has a calculated $pK_a$ of 10.9 and is thus also strongly shifted towards a more basic value. We conclude that Tyr41 is the proton donor in *Cr*LOV1-F41Y/C57A and C57S. In the case of *Cr*LOV1-F41Y, however, we can not exclude Cys57 as a possible proton donor. It should be noted that the accessibility of the solvent molecules into the binding pocket should be considered within the context of protonation. This will be discussed in the MD section.

### 4.3.4 Decay of the $S_1$ state suggests electron transfer via the excited singlet state

The TA data of the *Cr*LOV1-F41Y mutants provide evidence for the occurrence of an electron transfer reaction. Next to the observation of FMN and Tyr radical species, the triplet decay times obtained are as short as in *Cr*LOV1 wt indicating a quenching of the excited FMN triplet state. The only exception is *Cr*LOV1-F41Y/C57G in which no radical species could be detected in the $\mu$s time range and the triplet decay is prolonged compared to the other F41Y mutants. Determination of the $S_1$ decay times (table 4.1) yields a biexponential decay behavior for all *Cr*LOV1-F41Y mutants, including *Cr*LOV1-F41Y/C57G. In all F41Y mutants one decay time corresponds to the expected value for *Cr*LOV1 (i.e. 3 ns for wt and 4.5-5 ns for photo-inactive *Cr*LOV1 mutants) while the second $S_1$ decay time is a component that is faster than the time resolution of our apparatus (ca. 1 ns). This indicates that the CrLOV1-F41Y mutants can adopt at least two conformations to account for the biexponential decay behavior.

Furthermore, the fluorescence quantum yields, $\Phi_F$, of the *Cr*LOV1-F41Y mutants were determined (table 4.1). The fluorescence of the photoinactive mutants C57S, C57A and C57G are strongly quenched from $\Phi_F = 0.30$, 0.30 and 0.25 to $\Phi_F = 0.18$ (F41Y/C57S), 0.15 (F41Y/C57A), and 0.10 (F41Y/C57G) when additionally introducing Tyr41. The wt and the F41Y mutant have similar values of 0.16 and 0.14, respectively. We conclude from our measurements of $\Phi_F$ and the $S_1$ decay times that both, the photoactive C57 and the extra tyrosine, provide quenching channels that do not exist in the photoinactive mutants. This could be enhanced ISC to the triplet state, but might also indicate that some fraction of the electron transfer occurs from the singlet state.

To check this, we analyzed the sum of all DADS of each particular F41Y mutant. The sum of all DADS corresponds to the initial situation after the laser excitation pulse (ca. 10 ns) has ended, and before any $\mu$s-processes occur. These will be named hereafter $t_0$ spectra. They are shown in figure 4.3.7A. All spectra were scaled to the same value at 715 nm. At this wavelength exclusively the FMN triplet state absorbs and no other species contribute to the spectrum. As a reference, the $t_0$ spectrum of *Cr*LOV1 C57G is included. It is known from a previous study that this spectrum refers to the pure triplet spectrum of FMN in the *Cr*LOV1 core [12]. When comparing this reference spectrum with the $t_0$ spectra of *Cr*LOV1-F41Y, F41Y/C57S and F41Y/C57A significant differences become apparent in the range of 500 nm to 660 nm and below 410 nm (see figure 4.3.7A). Subtraction of the C57G reference spectrum from the $t_0$ spectra of *Cr*LOV1-F41Y, F41Y/C57S and F41Y/C57A yields the spectra in figure 4.3.7B, which show characteristic features of the flavin radical spectra described above. Since these spectra can already be observed at $t_0$, the radical species in *Cr*LOV1-F41Y, *Cr*LOV1-F41Y/C57S and *Cr*LOV1-F41Y/C57A can not be formed exclusively via a triplet state reaction. Thus, the electron transfer in these mutants must occur to some extent in the excited singlet state of FMN.

The $t_0$ spectrum of *Cr*LOV1-F41Y/C57G, on the other hand, is very similar to that of *Cr*LOV1-C57G except for the spectral range of the ground state bleach. This means that *Cr*LOV1-F41Y/C57G performs an electron transfer via the excited singlet state, as suggested by the fluorescence data. The radical is not stabilized and recombines on a timescale faster than formation of the triplet state.

**Table 4.1** Fluorescence quantum yields, $\Phi_F$, and fluorescence decay times of all F41Y mutants.

| | $\tau_1$/ns | $\tau_2$/ns | $\Phi_f$ |
|---|---|---|---|
| **F41Y** | 3.0 ($\sim$50%) | 0.9 ($\sim$50%) | 0.14 |
| **WT**[a] | 2.9 | - | 0.16 |
| **F41Y/C57A** | 3.4 ($\sim$50%) | 1.2 ($\sim$50%) | 0.15 |
| **C57A** | 4.5 | - | 0.30 |
| **F41Y/C57S** | 3.6 ($\sim$65%) | 1.0 ($\sim$35%) | 0.18 |
| **C57S**[a] | 4.6 | - | 0.30 |
| **F41Y/C57G** | 4.3 ($\sim$80%) | 0.9 ($\sim$20%) | 0.10 |
| **C57G** | 4.5 | - | 0.25 |

[a] Values for CrLOV1 wt and CrLOV1-C57S are taken from Holzer *et al.* [115].

### 4.3.5 Molecular Dynamics Simulations

The microenvironment surrounding the FMN chromophore and its interactions with the protein is of crucial importance. The photoreactivity of riboflavin is strongly affected by its hydrogen bond network [106, 116]. We performed 20 ns MD simulations of *Cr*LOV1-F41Y using the dark-state crystal structure of *Cr*LOV1 wt (pdb 1N9L) as a starting structure [58]. The MD data reveal that substitution of Phe41 with Tyr has only minor influence on the overall structure of *Cr*LOV1-F41Y. Small perturbations occur around Tyr41 but the hydrogen bonding network built around FMN C2(=O), N3(H) and C4(=O) including Asn89 and Asn99 remains intact. Tyr41 can form a hydrogen bond with FMN C4(=O) as well as with FMN N5 instead of Gln120, which is rotated away from FMN. A close-up view of the binding pocket of the 20 ns simulation structure of *Cr*LOV1-F41Y in comparison with the dark state crystal structure of the wild type is shown in figure 4.3.8. Figure 4.3.9A and B show the structures of *Cr*LOV1-F41Y/C57S and *Cr*LOV1-F41Y/C57A. In *Cr*LOV1-F41Y/C57S the hydrogen bond network between Asn89, Asn99 and FMN C2(=O), N3(H) and C4(=O) remains intact. In this case, the distance between Tyr41 and FMN C4(=O)

is small enough to form a hydrogen bond, while direct interaction with FMN N5 like in *Cr*LOV1-F41Y is quite unlikely. Instead, FMN N5 can form a hydrogen bond with the polar side chain of Ser57. In the case of CrLOV1-F41Y/C57A (figure 4.3.9B) the hydrogen bond network on the polar side of FMN is changed and does not include Asn99. Asn89 in turn, can form a hydrogen bond with N3(H) and C4(=O) of FMN.

The introduced Tyr41 is very close to the FMN and well stabilized in this mutant, with possible hydrogen bonds to FMN N5, C4(=O) and Asn89. Since Ala57 cannot form hydrogen bonds, Tyr41 is the only amino acid in the close proximity of FMN which can make a connection with FMN N5. In both mutants, *Cr*LOV1-F41Y/C57S and *Cr*LOV1-F41Y/C57A, Gln120 is rotated away from FMN, not interacting with the chromophore.

In contrast, FMN N5 does not participate in hydrogen bond interactions in *Cr*LOV1-F41Y/C57G (figure 4.3.9C). The distance between Tyr41 and FMN C4(=O) and FMN N5 is approx. 5 Å and therefore larger than in the other F41Y mutants. Furthermore, the hydrogen bond network between FMN and Asn89 is changed compared to *Cr*LOV1 wt. Asn89 can interact with FMN N3(H) and C4(=O), but due to a rotation of Asn99 no interaction between this asparagine and FMN is possible. Gln120 is interacting with C4(=O) instead.

Figure 4.3.9D and E compares the distances between Tyr41-OH and either FMN N5 (D) or FMN O4 (E) of the distinct F41Y mutants. The distances between Tyr41 and FMN N5 vary strongly between the different mutants from around 3.0 Å for *Cr*LOV1-F41Y to 5.0 Å for *Cr*LOV1-F41Y/C57G. The distances of *Cr*LOV1-F41Y/C57A and C57S are broadly distributed with a maximum at 3.0-4.0 Å. The distances between Tyr41 OH and FMN O4 on the other hand are almost equal in all F41Y mutants between 2.7 and 3.0 Å (figure 4.3.9E).

The structural differences of the particular CrLOV1-F41Y are accompanied by varying accessibilities for solvent molecules into the binding pocket. In *Cr*LOV1-F41Y, a slow water exchange was observed, with always one water molecule being present in the binding pocket. A very fast water exchange occurs in *Cr*LOV1-F41Y/C57G with constantly three to four water molecules present. *Cr*LOV1-F41Y/C57A also shows fast water exchange. The water molecule is located around FMN-O4, but the pathway for the HOH entry is different. While in the other *Cr*LOV1-F41Y mutants the water enters via the phosphate backbone of FMN, in *Cr*LOV1-F41Y/C57A a new water channel opens on the polar side of FMN. Solely in *Cr*LOV1-F41Y/C57S, no water exchange was observed. During the whole simulation time one non-exchanging water molecule is present, located between Ser57 and FMN oriented towards N10.

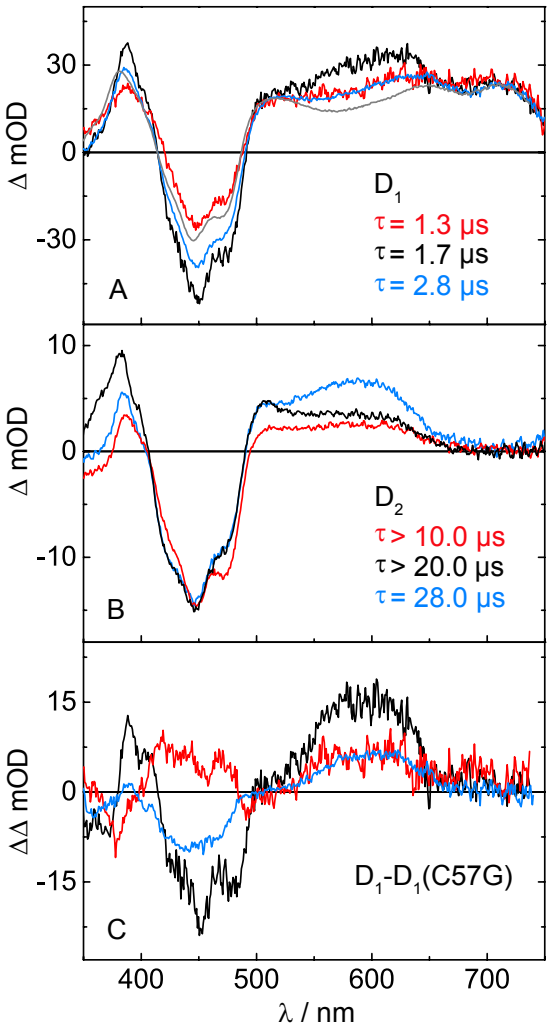


**Figure 4.3.5** DADS obtained from the global lifetime analysis for *Cr*LOV1-F41Y/C57S (blue lines) and *Cr*LOV1-F41Y/C57A (black lines). The DADS in **A** decay with a time constant of 2.8 $\mu$s and 1.7 $\mu$s. **B:** The second DADS of *Cr*LOV1-F41Y/C57S decays with a time constant of 28.0 $\mu$s, while the one of *Cr*LOV1-F41Y/C57A is non-decaying within the measured streak window of 20 $\mu$s. The DADS were scaled to the same minimum at 447 nm. The DADS of *Cr*LOV1-F41Y (red lines) are shown for comparison. **C:** The $D_1$ of panel A after subtraction of the $D_1$ of the C57G mutant (gray line in **A**), which is a pure triplet transient absorption.

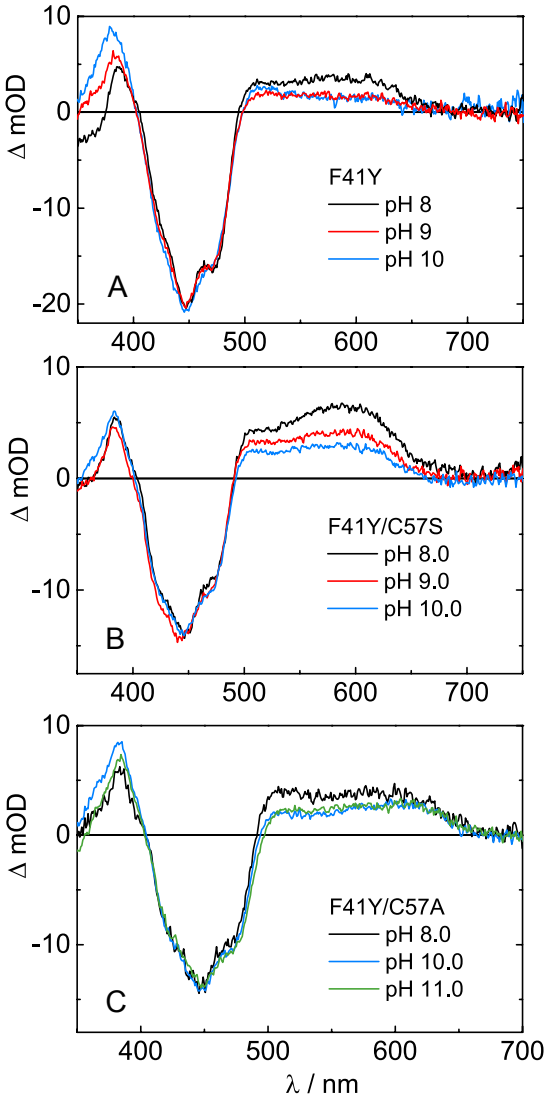


**Figure 4.3.6** Comparison of the $D_2$ of *Cr*LOV1-F41Y (**A**), F41Y/C57S (**B**) and F41Y/C57A (**C**) at pH 8.0, 9.0, 10.0 and 11.0, respectively.

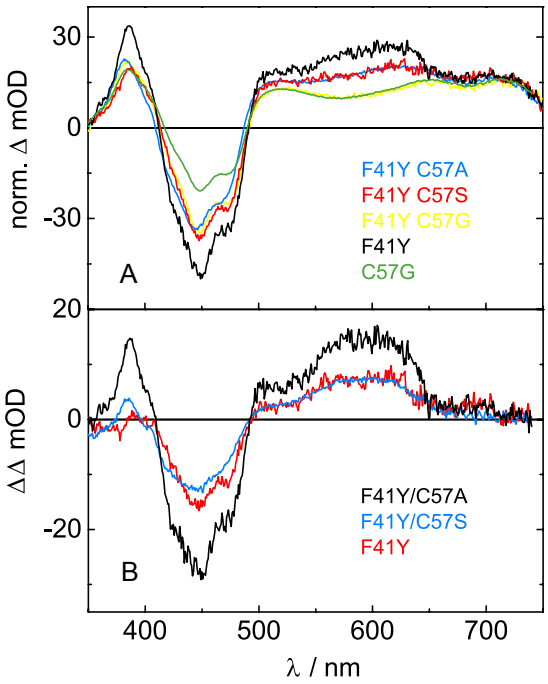


**Figure 4.3.7 A:** The $t_0$ spectra, i.e. the sum of all DADS, of *Cr*LOV1-F41Y, *Cr*LOV1-F41Y/C57S, *Cr*LOV1-F41Y/C57G and *Cr*LOV1-F41Y/C57A. The corresponding spectrum of *Cr*LOV1-C57G representing the pure triplet spectrum of FMN in the *Cr*LOV1 core is shown for comparison. **B:** Double difference spectra obtained by subtracting the $t_0$ spectrum of *Cr*LOV1-C57G from the $t_0$ spectrum of each F41Y mutant in **A**. These double difference spectra show characteristic features of FMN radical species.

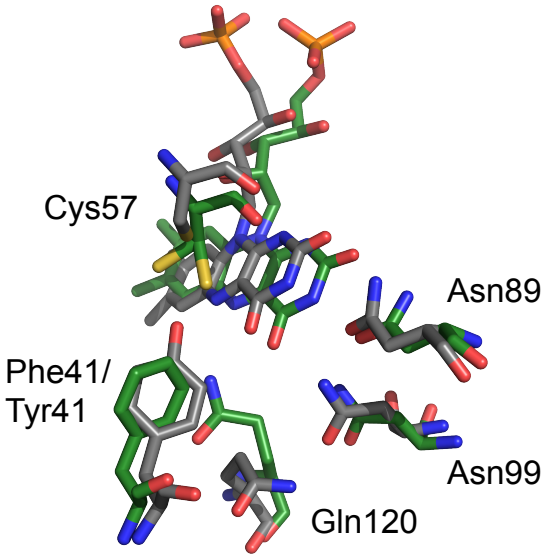


**Figure 4.3.8** Superposition of the 20 ns MD simulation structure of *Cr*LOV1-F41Y (gray) with the dark state crystal structure of *Cr*LOV1 wt (green) visualized by using PyMOL [117]. The mutation F41Y does not cause significant perturbations of the *Cr*LOV1 structure except for Gln120, which is rotated away from FMN.

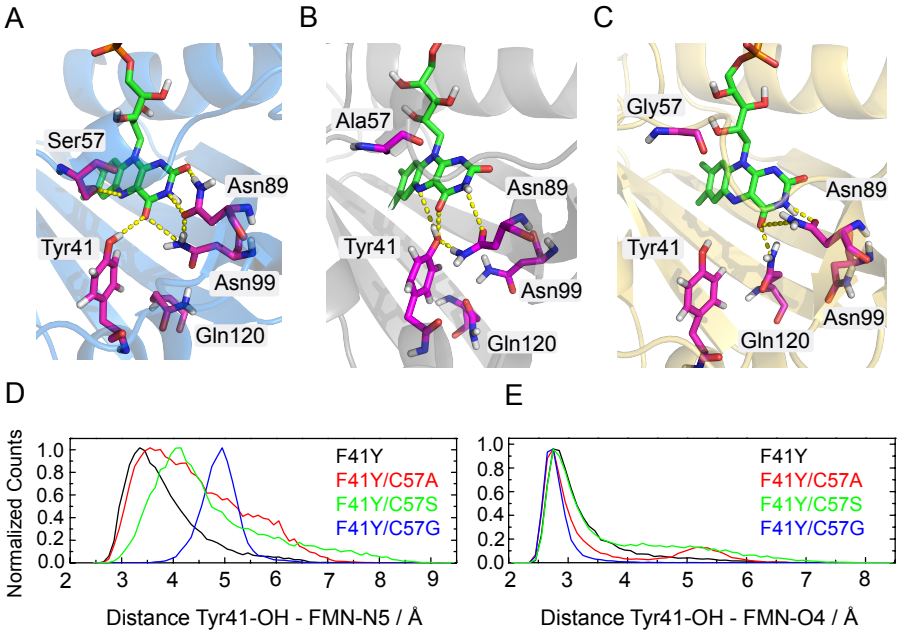


**Figure 4.3.9** 20 ns simulation structures of **A:** *Cr*LOV1-F41Y/C57S, **B:** *Cr*LOV1-F41Y/C57A and **C:** *Cr*LOV1-F41Y/C57G. Possible hydrogen bonds are indicated by dashed yellow lines. **D and E:** Distance distribution plots of F41Y mutants between Tyr41-OH and FMN-N5 and O4, respectively.

## 4.4 Discussion

Adduct formation as the initial step of activation is a well known attribute of LOV domains. While the reaction mechanisms of other flavin based blue light sensors, e.g. BLUF and CRY, implicate electron transfer reactions, no other light induced photoreaction than adduct formation has been observed so far in the case of LOV domains if the reactive Cys was present. In this study, we present spectroscopic and simulation data on *Cr*LOV1 mutants, in which a Tyr was introduced as a potential electron donor in close proximity to the flavin chromophore. The substitution of Phe41 by Tyr leads to a change in the primary photoreaction of *Cr*LOV1. The naturally preferred adduct formation reaction is inhibited and instead an electron transfer can be observed.

The mutation did not significantly alter the electronic properties of FMN since the ground state absorption spectrum shows only minor changes in comparison to *Cr*LOV1 wt. Moreover, MD simulations of both, *Cr*LOV1 wt and *Cr*LOV1 F41Y, reveal that both structures are highly similar with mainly preserved hydrogen bond interactions between FMN and the surrounding amino acids. Only minor perturbations around the introduced Tyr occur.

TA measurements in the $\mu$s time range showed that *Cr*LOV1-F41Y populates the excited FMN triplet state like observed for the wt, which decays also with a similar rate constant. But instead of the characteristic flavin-cysteinyl-adduct, radical species of FMN evolved from the triplet state. Spectral fitting of the data demonstrated the presence of $FMNH^{\bullet}$, $FMN^{\bullet-}$ and $TyrO^{\bullet}$. Likewise, when additionally substituting the reactive Cys57 with Ala or Ser, the formation of flavin radical species can be observed. *Cr*LOV1-F41Y/C57A and C57S as well populate the FMN triplet state that decays with comparably short lifetimes like observed for *Cr*LOV1-F41Y and wt. However, in these cases the deactivation channel of the triplet is electron transfer. Other photo-inactive *Cr*LOV1 mutants, in which electron transfer occurs only in the presence of an external electron donor like EDTA, usually have a triplet decay time by at least a factor of ten longer [4, 12]. Presumably, Tyr41 is the initial electron donor in the F41Y mutants although the amino acid sequence of *Cr*LOV1 wt contains four Tyr and one Trp as natural candidates. According to the crystal structure, they are located quite distant from the acceptor FMN, with Tyr47 being closest with a distance of 11.0 Å between Tyr47-OH and FMN-O4. Tyr41 on the other hand, is located very close with an average distance of 2.75 Å to FMN-O4 in each F41Y mutant providing a favored initial situation since electron transfer reactions are strongly distance dependent processes. Furthermore, the described electron transfer reactions are only observed, when Tyr41 is present.

On the other hand, our findings on the excited singlet state deactivation, including a reduced fluorescence quantum yield and a biexponential fluorescence decay, lead to the hypothesis that a second pathway for electron transfer via the excited singlet state exists in the *Cr*LOV1-F41Y mutants. This is supported by the fact that the

radical species are already formed at $t_0$ in the TA measurements. As another reaction proceeding from the excited singlet state of FMN, adduct formation in the LOV2 domain from *C. reinhardtii* (*Cr*LOV2) has been reported by Zhu and co-workers [118]. This was concluded from the ratio of transient absorption of the FMN triplet and transient bleach of the FMN ground state after decay of the excited singlet state. This ratio is smaller for *Cr*LOV2 than for *As*LOV1 (*Avena sativa* LOV1), indicating that some fraction of the excited singlet decays to a different species than the triplet state or the ground state. A deuterium effect observed in the fluorescence decay is taken as evidence that the primary step of this decay involves proton transfer to the excited singlet state of FMN. Interestingly, another study [12] found exactly the same ratio of both transients in *Cr*LOV1 and *Cr*LOV2. The latter study used LOV constructs with a His-tag, which are known to exist entirely as dimers or larger aggregates [99] whereas Zhu *et al.* [118] used LOV constructs fused to a maltose binding protein.
A reasonable explanation for the two reaction pathways of the F41Y mutants would be the existence of more than one conformation of each F41Y mutant most likely regarding Tyr41. Using a mutational approach on the DASH-type cryptochrome CRYD from *Synechocystis*, Biskup and coworkers have demonstrated that changes in geometry can change electron transfer pathways in CRYD and that distance is not necessarily the decisive parameter for electron transfer [119]. A theoretical study reported that already subtle conformational changes of the terminal electron donor Trp324 can inhibit the electron transfer along the Trp triad in *A. thaliana* cryptochrome [120]. The authors explain this by a drastic change in the electronic coupling between the donor Trp and the acceptor FAD. In our F41Y mutants we observe a broad dihedral angle distribution including FMN and Tyr41 (see supplementary, figure S4†) indicating a variation in the donor to acceptor orientation. Since *Cr*LOV1-F41Y, F41Y/C57S and F41Y/C57A feature the same photo-chemical characteristics upon blue-light excitation we propose that these mutants follow the same mechanism, meaning that the reactive Cys in *Cr*LOV1-F41Y apparently becomes unreactive. Although the orientation and distance of Cys57 and FMN seem to be favorable for adduct formation according to our MD simulations, we do not observe any matching signals in our spectroscopic data. The triplet decay time of *Cr*LOV1-F41Y is 1.3 $\mu$s and therefore in the same time range as in the wt. Consequently, the first step leading to the radical pair in the mutants containing Tyr41 is about as fast as the first step leading to the adduct in the wt. In the F41Y mutants this first step should be electron transfer from Tyr to $^3$FMN, for the wt the first step most likely is also electron transfer, this time from Cys57 to $^3$FMN. Accordingly, in F41Y we should expect both, adduct and FMN radicals, with similar yield.

In a previous study of the *Cr*LOV1 wt photocycle we could not detect any intermediate between the FMN triplet state and the adduct [12]. We concluded that the reaction is either concerted or it occurs as a sequence of electron transfer, proton transfer and radical recombination, with electron transfer being the rate limiting step. If adduct formation is a concerted reaction, it should compete with electron

transfer from Tyr in the F41Y mutant, and both products should be observed. If however, the reaction occurs in several steps, competition between the two reactions might occur only in the second step, meaning that after electron transfer from Cys the formed cysteine radical cation reacts faster with the Tyr than with the FMN radical anion. This is schematically shown in Figure 4.4.1. Here, $k_C$ and $k_T$ denote the rate constants for electron transfer from Cys and Tyr, respectively, which are both of similar magnitude of ca. $10^6$ $s^{-1}$. In the wt, adduct formation occurs then with a rate constant $k_A \gg k_C$. If we assume that in the F41Y mutant the cysteine radical cation oxidizes the Tyr with a rate constant $k_X \gg k_A$, no adduct should be formed. Oxidation of Tyr is always associated with proton transfer since the $pK_a$ value of tyrosine drops from 10.1 to -2 leading to inevitable deprotonation of $TyrOH^{\bullet+}$ to avoid high-energy pathways [121, 122]. If this deprotonation is faster than $k_C$ and $k_T$, $TyrOH^{\bullet+}$ can not be observed, and the products are the same as in the mutants lacking the reactive Cys.

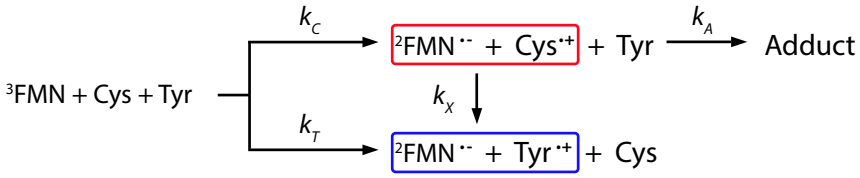


**Figure 4.4.1** Scheme of the proposed triplet state reaction mechanism of *Cr*LOV1-F41Y.

*Cr*LOV1-F41Y/C57G is the only mutant investigated which did not show a radical contribution in the $\mu$s time range. The excited triplet state is populated in this mutant but deactivates non-radiatively to the ground state with a biphasic decay by a factor of ten to 100 slower compared to the other F41Y mutants. Therefore, it behaves like *Cr*LOV1-C57G, i.e. the F41Y mutation does not influence the triplet state reaction in *Cr*LOV1-F41Y/C57G [12]. A similar behavior has been reported for a LOV based reporter protein in which a two-conformer model with different degrees of oxygen accessibility and therefore oxygen quenching has been suggested as an explanation [112]. This in turn indicates a very exposed flavin which would be accessible not only for oxygen but also for solvent molecules. And in fact, in *Cr*LOV1-F41Y/C57G we observed a very fast water exchange in the binding pocket.

Nevertheless, a strong quenching of the singlet state lowering the fluorescence quantum yield and lifetime can be observed which we consider to be a consequence of electron transfer in the excited singlet state. It is known from cryptochromes that protonation of FMN N5 by a nearby Asp supports the formation of a stabilized radical pair in flavoproteins. E.g., in the case of *A. thaliana* cryptochrome, proton transfer from D396 to FAD was reported resulting in the very stable radical ion pair $FAD^{\bullet-}$+$W(H)^{\bullet+}$ [123, 124]. In contrast, Langenbacher *et al.* [125] as well as Thöing *et al.* [126] observed proton transfer after electronic excitation of flavin in plant cryptochrome PHR from *Chlamydomonas*, i.e. formation of the neutral radical $FADH^{\bullet}$. Müller *et al.* observed both alternatives for *A. thaliana* cryptochrome depending on

pH and ATP binding [127]. According to our MD simulations the distance between Tyr41 and FMN is too large for direct protonation on very fast timescales and no alternative proton donor is located nearby FMN N5 in *Cr*LOV1-F41Y/C57G. Therefore, we propose that radicals formed in F41Y-C57G in the excited singlet state are not stabilized and recombine on a sub-$\mu$s timescale. The fact that two different reaction pathways are supported in *Cr*LOV1-F41Y/C57G leads to the conclusion that the FMN reactivity in this mutant is as well geometry-dependent.

The behavior of *Cr*LOV1-F41Y/C57G indicates that protonation of FMN N5 is significant for the stabilization of the radical pair $FMNH^{\bullet}$ $^{\bullet}$OTyr. Following the hypothesis of a PCET reaction from Tyr41 towards FMN, the electron as well as the proton donor is Tyr41 in *Cr*LOV1-F41Y, *Cr*LOV1-F41Y/C57S, and C57A. This is further supported by our TA measurements at basic pH values. The protonated $FMNH^{\bullet}$ could still be observed even at strongly basic pH of the solvent. This proves that protonation of the flavin radical can not be accomplished by water molecules in the binding pocket and that the proton donor must have a $pK_a$ value higher than 11. With a calculated pKa value of $> 12.00$ in *Cr*LOV1-F41Y, F41Y/C57A and 11.45 in F41Y/C57S, Tyr41 is an appropriate candidate for the required pH range. Moreover, structures of the mutants derived from MD simulations reveal that Tyr41 is the only proton donor present in the binding pocket besides Cys57, which is only present in *Cr*LOV1-F41Y. Our MD data further provide information about the accessibility of bulk water molecules into the binding pocket of the particular F41Y mutants. In any case, only one water molecule is located near the FMN at a given time in every particular mutant. If, however, this would be sufficient to deprotonate FMN N5 of the formed $FMNH^{\bullet}$, which we consider highly unlikely, the $pK_a$ value of this flavin neutral radical must also be comparably high. The $pK_a$ of $FMNH^{\bullet}$ in solution is 7.5, but can change drastically when incorporated in a protein [128].

## 4.5 Conclusions

Adduct formation as the primary step of activation is highly conserved among LOV domains. Here, we show that the reaction type can be switched to electron transfer by replacing a single amino acid that is not involved in the original reaction. Our findings not only provide implications for the reaction mechanism of the wild type adduct formation, it may also lead to the design of new artificial photosensors or photoswitches that are useful as optogenetic tools. Therefore, the applicability of the F41Y mutation for other LOV systems needs to be explored in the future.

# 5 Predicting the complete electron transfer pathways in proteins

## 5.1 Introduction

Electron transfer (ET) proteins control a large number of cellular functions and biological transformations, such as respiration and photosynthesis [129, 130] in various organisms. The role of the protein environment between the donor and acceptor sites and the effects it has on ET are still not fully understood. [131, 132] Understanding the complete ET pathway is thus of key importance in the study of ET proteins. The theoretical framework to understand electron transfer processes has been previously fully described [19–21, 133–135]. Marcus theory predicts that ET rates are determined by two terms, the Franck-Condon factor, which accounts for the thermal nuclear motions that bring the donor and acceptor into energetic degeneracy, and the electronic coupling between the states [27, 131]. For proteins, the electronic coupling is small due to the large distance between donor-acceptor pairs. Hence, the long-range electron tunneling rate constant can be derived from the Marcus formula [20, 21, 133], as shown in equation 5.1

$$k_{ET} = \frac{2\pi}{h}|V_{DA}|^2 \frac{1}{\sqrt{4\pi\lambda k_B T}} exp[-\frac{(\Delta G^0 + \lambda)^2}{4\lambda k_B T}] \tag{5.1}$$

where $V_{DA}$ is the electronic coupling matrix element between the reactant and the product electronic wave functions, $\Delta G^0$ is the Gibbs free energy change, and $\lambda$ is the reorganizational energy required to adjust the geometry of the system upon ET.

A simple empirical approximation, shown in equation 5.2, with a Gaussian form of Marcus equation, has been previously described by Moser and Dutton [19]

$$\log k_{ET} = 13 - \frac{\beta}{2.303}(R - 3.6) - \gamma\frac{(\Delta G^0 + \lambda)^2}{\lambda} \tag{5.2}$$

where $k_{ET}$ is the electron transfer rate in units of $s^{-1}$, $R$ (in Å) is the edge-to-edge electron transfer distance, $\Delta G^0$ (in eV) is the Gibbs free energy change and $\lambda$ (in eV) is the reorganizational energy. The coefficient $\gamma$ includes room temperature constants, as suggested by previous studies in photosynthetic reaction centers [19, 136, 137]. The parameter $\gamma$ has a value of 3.1 which comes from an empirical fit

of the Marcus plot that has broader inverted parabolas than in the classical fit [138]. The reorganizational parameter $\lambda$ has a typical value of $0.9 \pm 0.2$ eV [139] in proteins, while the parameter $\beta$, which describes the penetration of the electronic wave functions in the redox centers' medium [19], has a value of 1.4 $Å^{-1}$ [137]. Using these values for $\beta$ and $\gamma$, the Moser-Dutton equation, in its originally published form [23], becomes the following equation (equation 5.3)

$$\log k_{ET} = 13 - 0.6(R - 3.6) - 3.1\frac{(\Delta G^0 + \lambda)^2}{\lambda} \tag{5.3}$$

Previously published works on calculating the ET rates in proteins are based on calculations of electronic structures or optical properties of simplified systems [22, 25, 140]. There have also been models developed that included the inhomogeneity of the protein medium [141]. Models that include the tunneling pathways between the donor and the acceptor by combining the through-space, covalent and hydrogen-bonded contacts have also been described [25, 142]. Most structure-based ET coupling theories require parameters that describe the electronic structures of the donor and acceptor states. However, in most cases, these parameters cannot be fully described. Hence it is often assumed that only one electronic wave function is involved [22]. QM/MM is often the method used to model ET pathways in proteins [143, 144]. However, QM/MM has a very high computational cost and takes long time to complete and analyze [145]. Additionally, QM/MM requires the user to have a significant background in computational chemistry to successfully setup, execute and analyze the calculations.

To bypass these issues and to ease the process when working with ET proteins, we have developed a novel set of algorithms and implemented them in etFIND. In addition to predicting the full ET pathway in proteins, etFIND is able to predict both the initial residue where the ET chain will originate from as well as the terminal acceptor residue where the ET chain will terminate at, neither of which has been previously reported in the literature. etFIND can also be used for *in silico* engineering of ET pathways in proteins and their mutants and is able to predict both intra- and inter-protein ET pathways, as discussed below.

## 5.2 Algorithms

### 5.2.1 Finding the optimal pathway

The ET-pathway-finder algorithm calculates the ET transfer rates based on the structural information of a protein and predicts the full ET chain within the given structure. The algorithm includes five major steps: (i) set the donor residue, (ii) calculate the rates between the donor residue and all other residues in the protein,

(iii) find the largest rate and set a new donor residue, (iv) repeat steps i-iii until the final acceptor of ET is reached and (v) find the optimal pathway.

To calculate the rates, ET-pathway-finder uses equation 5.3 in combination with Novotny's free energy formulation [146, 147], where $\Delta G$ is defined as a difference in free energies of states before and after an electron hopping has taken place.

$\Delta G$ is given [146] as:

$$\Delta G = \Delta G_{\phi} + \Delta G_{EL} - T\Delta S_{CF} - T\Delta S_{TR} - T\Delta S_{CR} \tag{5.4}$$

where $\Delta G_{\phi}$ describes hydrophobic effects while $\Delta G_{EL}$ describes electrostatic interactions and hydrogen bonding. Conformational entropy is accounted for by $T\Delta S_{CF}$, while $T\Delta S_{TR}$ describes translational and rotational entropy and $T\Delta S_{CR}$ includes cratic entropy effects (entropy change corresponding to infinite dilution). The structures used in our calculations are static with no structural changes before and after an electron hopping. As a consequence all terms in $\Delta G$ cancel out, except the electrostatic contribution.

The Coulomb potential used to evaluate all atomic pair-wise interactions is shown in equation 5.5, where the sum is taken over the two states (before and after ET has taken place) of the given atom pair, for all possible atom pairs.

$$\Delta G = \sum_{i<j}^{n} \frac{q_i^{(a)} q_j^{(a)} - q_i^{(b)} q_j^{(b)}}{16\pi r_{ij}^2} \tag{5.5}$$

where $q_i^{(a)}$ and $q_i^{(b)}$ are atomic charges before and after the electron transfer, and $r_{ij}$ are inter-atomic distances.

The ET pathway finder algorithm makes no attempt to correct for the temperature-dependent Boltzmann factor [23], as it has been shown that this is unnecessary and may lead to inaccurate results [138]. Due to the high parametrization of the rate constant equation used, calculation of the full pathway is very fast and needs up to a few minutes to complete on a desktop computer (see Table 5.2).

The optimal pathway between the initial donor and the terminal acceptor is found by analyzing the calculated rates between all residues within a sphere around the center of mass of the protein. The radius of this sphere can be specified by the user, where the default value is the whole protein. This algorithm allows for a global search for an ET chain through the protein or protein pairs. This search is possible since the equations used for calculating the ET rates are highly parameterized, and approximations are used to balance execution times versus the accuracy.

### 5.2.2 Predicting the start and terminal residue

The algorithm for predicting the initial donor residue analyzes the structural information of the protein and calculates the rate for each potential donor residue being the start of the ET chain, and selects the residue with the highest rate. The calculations of rates are performed based on the equation 5.3, as previously described.

More specifically, the ET-centers-prediction module places a dummy acceptor residue at the protein's center of mass. These coordinates are used for the edge-to-edge distance calculations between each potential donor residue and the dummy acceptor residue. The dummy acceptor residue is used here as a terminal acceptor, as in certain cases the terminal acceptor residue is not present in the crystal structure. With this choice of a dummy acceptor it is not needed to model the true acceptor before the calculation is performed. For each tryptophan and tyrosine residue as well as ions present in the protein structure, a rate is calculated between that residue and the dummy acceptor residue. The residue with the highest probability is selected as the start of ET chain. Currently the prediction algorithm supports Fe, Cu, Zn, Co, Mg, Mn, Ni and Re ions.

To verify the methodology used for predicting the start residue, we have tested all possible start positions in each test case. For all possible starting positions we have calculated both the forward and backward optimal pathway and the total rate of the pathway, between initial donor and terminal acceptor residues. The efficiency $\phi$ of a particular path $a \to b \to c \to \cdots \to y \to z$ is calculated as

$$\phi = \frac{k_{ab}}{k_{ba}} \times \frac{k_{bc}}{k_{cb}} \times \cdots \times \frac{k_{yz}}{k_{zy}} \tag{5.6}$$

where $k_{i<j}$ is the rate of the forward step, while $k_{j>i}$ is the rate of the backward step. For a given terminal acceptor, the initial donor with the largest value of $\phi$ should be the most likely starting donor. In each of the tested proteins, we find the highest efficiency for the same residues that have been determined as start residues by the previously described simpler methodology, which considers only the direct rate between the start and the end donor. In all cases that we tested, these are also the same residues that were determined experimentally as the start of the ET chains.

This shows that it is not necessary to perform the elaborate and time-consuming calculations for all possible paths (as in eq. 5.6). Instead, it is sufficient to calculate the direct transfer rate to a single reference residue placed at the center of mass of the protein in order to find the initial donor residue.

The algorithm for predicting the terminal acceptor residue performs a search in a similar manner as the algorithm for predicting the initial donor residue. The only difference is that the dummy residue is now used as an electron donor center. The residue with the highest calculated ET rate is selected as a terminal acceptor. Ion screening is also included and supports the same ions as previously described.

### 5.2.3 Implementation of the algorithms

These algorithms have been implemented in the program package etFIND, which is written in a completely modular fashion. The fully modular nature of etFIND allows for users to add modules or include third-party programs to be executed in a sequence with their own modules, e.g. in the cases of data forwarding and output dependency.

As input, etFIND requires a pdbqt file and residue numbers of the initial donor and the terminal acceptor residues. The residue numbers of the initial donor and the terminal acceptor can both be predicted by the etFIND's modules, so the pre-knowledge of the system is not required. The output of etFIND is a string with residues that are found to be important in ET pathway.

## 5.3 Validation of algorithms

We have tested the described algorithms on previously published ET pathways in proteins and found that the algorithms are able to reproduce all the tested pathways as well as both the initial donor and terminal acceptor residues for each test case. The validation has been carried out on following systems: RNR R2 [148], DNA photolyase [149], cryptochrome 1 [124, 150], Methylamine dehydrogenase-Amicyanin complex [151], cytochrome c peroxidase-cytochrome c complex [129], both the wild type and the 122W mutant of azurin [152], cytochrome P450 BM-3 [153], metmyoglobin [154] and AppA BLUF [155]. These proteins have been selected based on previously published descriptions of their full ET pathways. In all tested cases the ET-pathway-finder algorithm predicted the same crucial residues to be involved in ET pathway and in the same pathway-order, as already described in the corresponding references. Comparison of the previously published pathways and the pathways predicted by the algorithm described here are shown in Table 5.1.

In the cases of DNA photolyase, metmyoglobin, AppA BLUF and cryptochrome-1 we find identical residues involved in the pathways like those already published for these proteins, without any additional residues involved. In other test cases, we identify a few additional residues that are likely to be involved as intermediates in the pathway, in addition to the crucial residues that have already been identified in the previous studies.

Additionally, we have validated the ET-centers-prediction algorithms (predicting both the initial donor and terminal acceptor residues) on the same systems and found that they predict the correct start and end of the ET chain in all of the tested cases. The results of these validations are shown in Table 5.1.

It is interesting to note here that all of our calculations have been performed on the crystal structures for that specific protein case. This suggests that it is not necessary to perform time-consuming structural optimizations or structural relaxations

in solution for these cases, as the results would not be affected. However, in certain cases it might still be necessary to perform structural optimizations for correct predictions of pathways and ET centers.

Table 5.2 lists timings for each test case. All calculations were performed on a single CPU core of a desktop computer with an Intel® Core™2 Quad-Core CPU Q9400 clocked at 2.66 GHz running a 64bit version of openSUSE® 11.2 (2.6.31.14-0.8-desktop x86_64) operating system. Meaning that the execution times shown here are also the total wall clock times. The time it took for the calculation of the pathway as well as the time it took for the prediction of the ET pathway point-of-origin residue as well as the terminal acceptor residues are included in the total execution times. The execution times vary from 16 seconds, in case of 2IYG with 919 atoms, up to 254 seconds for the 1U3C system which has 4022 atoms. As can be seen from Table 5.2, for these test cases, the algorithm scales linearly with the number of atoms present in the protein structure.

**Table 5.1** Results of the validation tests of ET pathways and ET centers prediction modules in comparison to the previously published pathways.

| Protein (structure) | Published pathway | Predicted pathway | Predicted donor | Predicted acceptor |
|---|---|---|---|---|
| DNA photolyase (1DNP [156]) | W306, W359, | W306, W359, W382, FAD | W306 W382, FAD | FAD |
| Methylamine dehydrogenase-Amicyanin complex (2GC4 [157]) | TRQ, H95, HOH, S56, Cu | TRQ, P52, P94, H95, S56, Cu | TRQ[a], Cu | Cu |
| Cytochrome P450 BM-3 (1BVY [153]) | a) FMN, P382, E387, C400, Hem b) FMN, P392, G393, R398, Hem | FMN, G396, Q397, A399, C400, I401, M490, G491, T492, Hem | FMN[a], Hem | Hem |
| Azurin (3IBO [158]) | Cu, Rep | Cu, M121, C112, Rep | Cu, Rep | Cu, Rep |
| RNR R2 (1MXR [159]) | Y122, H118, W48 | Y122, H118, W48 | Y122 | W48 |
| Metmyoglobin (1YMB [160]) | W14, Hem | W14, Hem | Hem | Hem |
| Azurin 122W (3IBO-122W[b]) | Cu, W122, Rep | Cu, M121, C112, W122, Rep | Cu, Rep | Cu, Rep |
| AppA BLUF (2IYG [161]) | Y21, FAD | Y21, FAD | Y21 | FAD |
| Cytochrome C peroxidase-Cytochrome C complex (2PCC [129]) | HemA, W191, G192, A193, A194, HemB | HemA, N196, N195, A194, A193, K179, G178, L177, A176, HemB | HemA, HemB | HemA, HemB |
| Cryptochrome 1 (1U3C [162]) | W324, W377, W400, FAD | W324, W377, W400, FAD | W324 | FAD |

[a] These residues have been identified as the start residues when the ion pre-screening has been disabled as a terminating step. [b] Structure generated using Modeller v9.14 software package [101] using 3IBO [158] crystal structure as a template for modeling.

**Table 5.2** Execution times on a single CPU core, including the ET pathways execution time as well as both ET centers prediction modules run times.

| Protein | Structure | Execution time (s) | Number of atoms |
|---|---|---|---|
| DNA photolyase | 1DNP | 184 | 3854 |
| Methylamine dehydrogenase-Amicyanin complex | 2GC4 | 59 | 1837 |
| Cytochrome P450 BM-3 | 1BVY | 211 | 4779 |
| Azurin | 3IBO | 24 | 997 |
| RNR R2 | 1MXR | 114 | 2783 |
| Metmyoglobin | 1YMB | 26 | 1242 |
| Azurin 122W | 3IBO-122W | 24 | 1003 |
| AppA BLUF | 2IYG | 16 | 919 |
| Cytochrome C peroxidase-Cytochrome C complex | 2PCC | 159 | 3304 |
| Cryptochrome 1 | 1U3C | 254 | 4022 |

## 5.4 Design of alternative pathways

DNA photolyases belong to the photolyase-cryptochrome superfamily. [163] These enzymes repair UV-induced DNA lesions, including cyclobutane pyrimidine dimers (CBT photolyases) and pyrimidine-pyrimidone photoproducts (6-4 photolyases). [164] CBT-photolyase binds to pyrimidine dimers and, upon blue light absorption, breaks the cyclobutene ring that is joining the two pyrimidines. DNA photolyase contains noncovalently bound fully reduced flavin adenine dinucleotide ($FADH^-$) as cofactor. The second cofactor required is either a 5,10-methenyltetrahydrofolate or a 5-deazariboflavin that acts as an antenna for blue-light absorption. So far, 12 elementary ET steps have been identified together with six ET reaction pairs. [165] The excited $FADH^\bullet$ acts as a sink that draws electrons from the surrounding aromatic residues of the entire protein. [165]

The mechanism of action of DNA photolyase has been previously extensively studied, however, the complete mechanism and repair dynamics have not yet been fully resolved. [162] We investigated the ET pathways of wild type DNA photolyase and of some of its mutants with our algorithms. Our aim is to further understand the ET within the protein and to determine the importance and the role of specific residues involved in the mechanism as well as involvement of alternative ET pathways.

We have used etFIND to analyze the pathway and guide the *in silico* mutagenesis of DNA Photolyase from *E. coli*. As a reference point for the wild type structure and as a modelling template we used an available crystal structure of DNA Photolyase, pdb code 1DNP [156]. The previously published mechanism [149] predicts

**Table 5.3** Pathways predicted by etFIND for each mutation in DNA photolyase.

| Mutation | Predicted pathway |
|---|---|
| Wildtype | W306, W359, W382, FAD |
| | W157, HOH606, FAD |
| | W277, FAD |
| | W316, W382, FAD |
| | Y365, FAD |
| W359G | W306, W382, FAD |
| W382G | W306, W359, FAD |
| W359G-W382G | W306, FAD |
| A348W-V352W-W359G-W382G | W306, W352, W348, FAD |
| A348W-W359G-W382G | W306, W348, FAD |
| V352W-W359G-W382G | W306, W352, FAD |

the dominant ET pathway as W306-W359-W382-FAD which is also the pathway predicted by etFIND. Based on this pathway, the following mutations were proposed and analyzed: W359G, W382G, W359G-W382G, A348W-V352W-W359G-W382G, A348W-W359G-W382G and V352W-W359G-W382G. Furthermore, we analyzed potential, alternative pathways in the wild type structure starting from different initial residues.

A recent study [166] proposed several potential ET pathways in DNA photolyase. The dominant pathway, known as the *classical pathway*, is reported to be W306-W359-W382-FAD. Additionally, two secondary pathways have been proposed: *alternative pathway*, Y365-W316-W382, and a hypothetical *five-Trp pathway*. The alternative pathway, Y365-W316-W382, shares the same W382 residue with the classical pathway, while the theoretical five-Trp pathway shares no common Trp residues with the classical nor with the proposed alternative pathways. It is important to note that the five-Trp pathway has been suggested for the PhrA structure only, while the other two pathways have been proposed to be involved in several proteins in addition to PhrA.

To generate each of the mutant structures, we have used the Modeller v9.14 software package [101] as well as the Swiss-Model web server [167] to verify the generated structures. Both software products predicted structures with minimal structural differences.

For each of the mutants listed above, etFIND finds a new dominant pathway. The predicted pathways for each of the mutants are shown in Table 5.3.

We tested several potential pathways. Starting from W306 in the wild type structure, etFIND predicted the classical pathway to be the dominant one. When the start residue is Y365, it predicts a direct transfer between Y365 and FAD. Together with the W316-W382-FAD secondary pathway, this suggests that the alternative pathway is less important than the classical pathway for the ET mechanism of wild

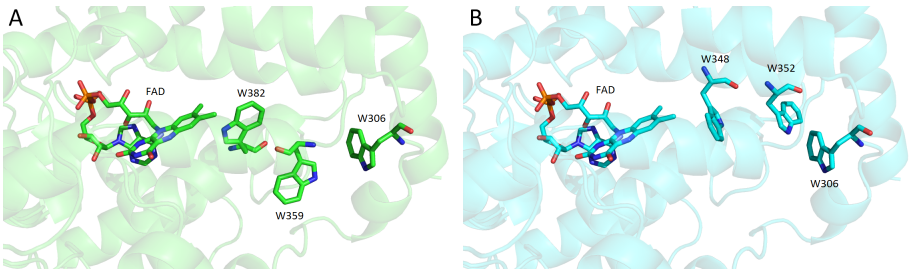


**Figure 5.4.1** Structural representation of predicted ET pathways in wild type (left, green) and A348W-V352W-W359G-W382G mutant (right, blue) of DNA Photolyase. Residues identified as important in the ET pathway of the respective structures, are shown as sticks. Images generated using the PyMOL molecular graphics system [117].

type DNA photolyase. Moreover, we find that the five-Trp pathway plays no role in the ET mechanism of wild type DNA Photolyase.

Furthermore, we propose that the quad-mutation, A348W-V352W-W359G-W382G has a completely different dominant ET pathway, when compared to the wild type. The dominant pathway is changed from W306-W359-W382-FAD in wild type to W306-W352-W348-FAD in the quad-mutant. The predicted pathways in the wild type and the quad mutant are shown in Figure 1. Interestingly, in the double mutant, W359G-W382G, etFIND does not identify any additional residues that would facilitate the ET. This suggests that in this double mutant no ET would occur or ET would not be a dominant mechanism of action. A thing to note here is that, while these mutations are predicted to have different modes of actions or a different ET pathway, it is uncertain whether these protein constructs can be experimentally expressed.

### 5.4.1 Predicting electron transfer pathway and donor and acceptor residues in CrLOV

We applied etFIND on the structures of CrLOV F41Y mutants. Structure of each mutant, in PDBQT file-format, was used to perform etFIND calculations. In all mutants etFIND predicts that Tyr41 is the donor residue, while the FMN is the acceptor, terminating the ET chain. For the ET pathway, etFIND predicts a direct transfer of the electron, from Tyr41 to FMN, without any intermediate residues involved as charge-carriers.

## 5.5 Conclusions

Here, we report a novel, efficient method for predicting residues involved in both the intra- and inter-protein electron transfer pathways. We also describe the algorithm for predicting the residue where the electron transfer will originate from within the

protein as well as the algorithm to determine the terminal acceptor residue of the ET chain.

The calculations of ET rates are based on the Moser-Dutton equation in combination with the free energy formulation of Novotny. In all of the test cases the predicted pathway was either identical to the previously published pathway for that protein or included additional residues that might be important in the ET chain, in addition to the crucial residues previously identified for that specific system. Due to the high parametrization of the equations used in calculating the ET rates in the global search for the pathway, the total execution times are very short. While the execution times depend on the total number of atoms present in the analyzed structure, the highest runtime (wall clock time) we have seen in our tests was 4 minutes and 14 seconds.

This algorithms successfully predicts the dominant ET chain for structural configurations with a single dominant ET pathway. In addition to the pathway algorithm, we report an algorithm that is able to predict the initial residue where the ET will originate from. We also propose an algorithm that is able to predict the terminal acceptor residue of the ET chain. Both algorithms have been validated on the same test cases as the ET pathway finder algorithm, and they are both able to predict the correct ET centers in all of the tested systems.

We further show the usefulness of the described method in guiding the *in silico* mutagenesis of DNA Photolyase. We propose that in the case of a A348W-V352W-W359G-W382G mutant there is a new dominant ET pathway, while in the case of W359G-W382G mutant no residues are identified as part of the ET chain which might suggest that this mutant may have a different mechanism of action, compared to the wild type. Our results for the wild type suggest that the classical pathway is the dominant ET mechanism, while the alternative pathway, Y365-W316-W382, contributes much less to electron transfers of DNA Photolyase. Additionally, we used the etFIND software package to estimate the donor and acceptor residues in CrLOV as well as the ET pathway between the two. In all CrLOV mutants etFIND predicts Tyr41 and FMN as the donor and acceptor residues, respectively, with direct ET between the two residues.

# 6 Signaling mechanism of Rac1 and Rac1b and their protein constructs with p67$^{phox}$, RhoGDI and LOV2

## 6.1 Introduction

The Ras superfamily of small GTPases controls a large number of biological functions and responses through specific bindings to effector proteins. [168] The Rho protein family is a subgroup of the Ras superfamily and is made of 22 mammalian members - three Rho isoforms A, B, and C; three Rac isoforms 1, 2, and 3; Cdc42, RhoD, Rnd1, Rnd2, RhoE/Rnd3, RhoG, TC10, and TCL; RhoH/TTF; Chp and Wrch-1; Rif, RhoBTB1, and 2; and Miro-1 and 2. [169–171] Rho proteins are guanine nucleotide binary switches that control signaling pathways by cycling through active and inactive forms, GTP-bound and GDP-bound states respectively, facilitated by guanine nucleotide exchange factors (GEF), GDP dissociation stimulators (GDS) and GTPase activating proteins (GAP). [168, 172–177] A schematic representation of a GTPase cycle is shown in figure 6.1.1. The most studied Rho proteins are Rac1, RhoA and Cdc42. [170] These have been shown to regulate gene expression, adhesion, migration, oxidase activity, actin assembly and motility, and apoptosis induction. [172, 178, 179] Several Rho family proteins have been shown to regulate oncogenic Ras transformation. [172]

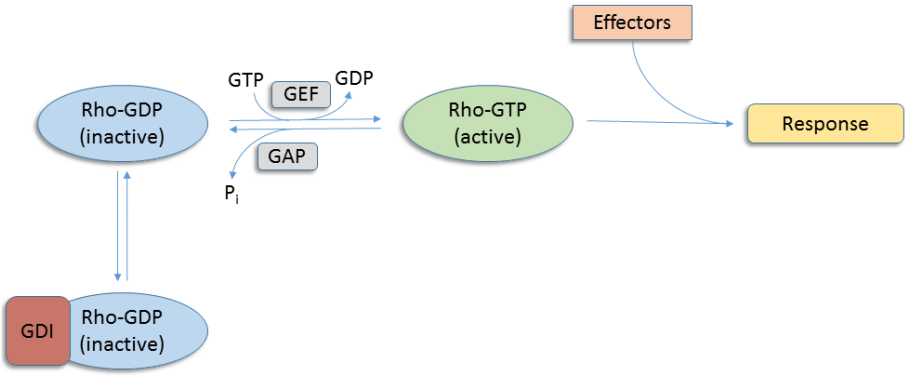


**Figure 6.1.1** Schematic representation of a GTPase cycle, showing the activation of a Rho protein and the induction of response upon the protein activation.

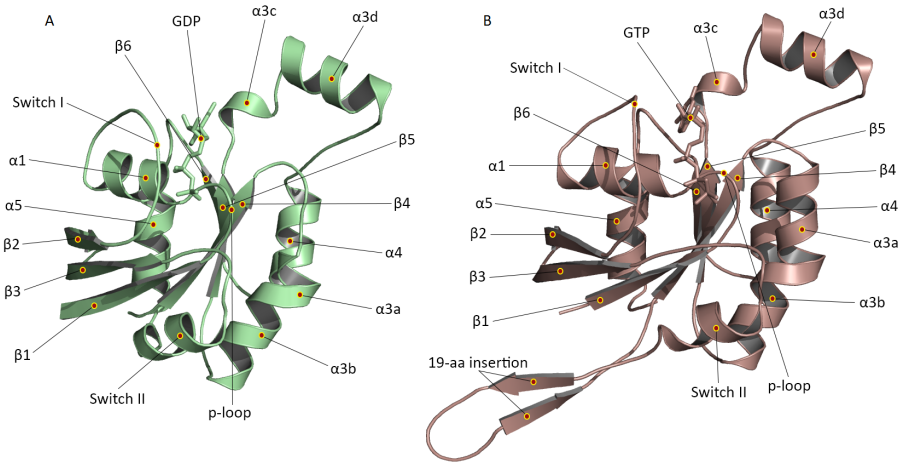


**Figure 6.1.2** Structural comparison of Rac1 (left, green) and Rac1b (right, dark-red) proteins. Key structural elements are labeled for each of the structures.

Rac1 is a member of the Rho protein family that cycles between an inactive GDP-bound and an active GTP-bound state. [180] Rac1 has been shown to regulate the vascular smooth muscle cell proliferation and migration and the flow induced re-alignment of endothelial cells. [181] Rac1 has an important role in cell motility and growth control, through formation of lamellipodia and membrane ruffles. [177, 182] It has also been linked with various other cellular processes such as phagocytosis, adhesion, cell proliferation and differentiation. [183–185] Rac1 has been implicated [186] in control of chondrogenesis through activation of the JAK/STAT3 signaling as well as control of activation of the MEKK-SEK-JNK kinase cascade through MLK3 binding. [187] Rac1 has been linked [188] with multiple renal and cardiac diseases. It has been shown [174] that activating-mutations of Rac1 cause tumorigenic transformation *in vitro*. However, it could not be shown *in vivo* that Rac1 contributes to malignancy in human tumors. However, de-regulated Rac activity contributes to various aspects of tumorigenicity. [189]

Rac1b is a spliced isoform of Rac1 that contains additional 19 residues inserted in the sequence of Rac1, close to the C-terminal. It is constitutively active with a very high guanine nucleotide cycling rate and decreased intrinsic GTPase activity. [174, 177] It is up-regulated in malignant breast, lung and colorectal cancers. [168, 174, 190, 191] There is also some evidence that Rac1b is involved in modulation of malignant progressions of thyroid [192, 193] as well as pancreatic [194] carcinomas.

Structural studies have shown that the switch 1 region, together with switch 2, is an important regulatory element of Rac1 and is crucial for effector interactions. [180] The p-loop is important for phosphate binding and magnesium coordination, in addition to the central region of switch 1 and the N-terminal part of switch 2. Hydrogen bonds between the guanine base and the switch 1 residues in Rac1 are water-bridged, there are no direct hydrogen bonds formed between these two regions, as opposed to the Rac1b. [175, 180] Switch 1 and 2 are the regions of the

highest structural differences between the Rac1 and Rac1b. In Rac1b, the switch 1 region is highly displaced from the nucleotide-binding site, and the switch 2 region is highly mobile compared to Rac. This leads to a decreased GTPase reaction of Rac1b and increased effector binding. Available experimental data of Rac1b show increased mobility of the Switch 1 and 2 regions, while the 19-amino acid insertion was not visible in the electron density map of its crystal structure. [175] In most of the GTPase crystal structures, the switch 2 region is highly ordered, indicating that the most likely cause for the switch 2 mobility is the very high mobility of the adjacent 19-amino acid insertion. [175] The switch 1 displacement, on the other hand, is most likely due to disrupted interactions between switches 1 and 2. This open conformation of switch 1 results in changed contacts of Thr35 with $\gamma$-phosphate and the $Mg^{2+}$ ion. This can explain the rapid nucleotide dissociation of Rac1b.[175] Figure 6.1.2 shows a side-by-side comparison of the Rac1 and Rac1b structures, with important structural elements labeled according to Rac1 and Rac1b numbering. [173]

While the amount of expressed Rac1b in cells is quite small, the fraction of activated Rac1b compared to that of activated Rac1 is often quite high. [195] This suggests that Rac1b has a crucial role in Rac downstream signaling in cells. It has been suggested that the 19 amino acid insertion is acting as a new effector-binding site in Rac1b, resulting in involvement in signaling pathways related to the normal or neoplastic growth of the intestinal mucosa. [196] Contrary to Rac1, Rac1b does not require GEFs to be activated and behaves as a self-activating GTPase. [175] GEFs and GAPs control the activities of both Rac1 and Rac1b, however the difference in activation of the two isoforms is controlled by the lack of interaction of Rac1b with Rho-GDI. As a consequence Rac1b remains bound to the plasma membrane and is not found in the cytoplasm. [195] Contrary to Rac1's regulation of lamellipodia formation, Rac1b shows no such control and is unable to bind-to and activate JNK and PAK protein families. Both proteins are able to activate the NF-$\kappa$B protein complex. However, only Rac1 but not Rac1b is able to further stimulate NF-$\kappa$B by controlling the RelB-NF-$\kappa$B2/p100 pathway. [197] This RelB channel has been reported to down-regulate the NF-$\kappa$B pathway, which might suggest that Rac1b is promoting tumorigenesis by activating NF-$\kappa$B signaling and bypassing the down-regulation through the RelB pathway. [197] A recent study [198] has shown that Rac1b promoted cell proliferation by activating the JNK2/c-JUN/cyclin-D1 pathway and inhibited apoptosis by modulating the AKT2/MCL1 pathway.

Reactive oxygen species (ROS) are oxygen-derived small molecules, including radicals, such as ozone, singlet oxygen, hydrogen peroxide, as well as superoxide, peroxyl and hydroxyl radicals. ROS production is usually a cascade of reactions which start with the production of superoxide. Superoxide easily converts to hydrogen peroxide either spontaneously or through superoxide dismutase catalysis. [199] As ROS highly and rapidly interact with other molecules, production of ROS is tightly controlled under physiological conditions.

Nicotinamide Adenine Dinucleotide Phosphate (NADPH) oxidases (NOXs) are important contributors to the generation of ROSs during hypoxia-reoxygenation. [200] NADPH oxidase is an enzyme complex that is a primary defense mechanism against microbial infections. [201] Mutation in any of the constituting subunits of NADPH oxidase can cause chronic granulomatous disease, due to issues in ROS generation. NADPH complex consists of six subunits: a Rho-GTP (usually Rac1 or Rac2), Cytochrome b-245 heavy chain (also known as $gp91^{phox}$ or Nox2), $p22^{phox}$, $p40^{phox}$, $p47^{phox}$ and $p67^{phox}$. [199] Cytochrome subunit of the NADPH oxidase has six homologs, which are collectively referred to as the NOX family. There are seven members in the NOX family: Nox1-Nox5, Duox1 and Duox2. [199]

In resting cells, $p40^{phox}$, $p47^{phox}$ and $p67^{phox}$ exist as a complex which, when activated, translocates to the membrane and binds Nox2 to form an active enzyme complex. Binding of Rac GTPase to $p67^{phox}$ is a crucial step in NADPH oxidase assembly and activation. While inactive, Rac protein (in GDP bound form) exists in a complex with RhoGDI (Rho GDP-dissociation inhibitor). Upon activation, this complex dissociates and Rac is translocated to the plasma membrane, independent of other components. [201]

$p67^{phox}$ binds to the cytoskeleton. In resting cells it forms an inactive complex together with $p40^{phox}$ and $p47^{phox}$, while in activated cells it interacts with Nox2 and Rac GTPase to form an activated enzyme. [202] $p67^{phox}$ also regulates the enzymatic activity of Nox2 by binding to it with an activation domain. [203]

Due to the important roles that Rho GTPases play, they are regulated on multiple levels. As described before, GEFs catalyze the GDP/GTP cycling, while GAPs accelerate inactivation and signal termination. A third level of regulation of Rho GTPases is by RhoGDI. RhoGDI interacts only with prenylated[1] Rho proteins. [204] The site for post-translational modifications is the conserved C-terminus CAAX sequence. [205] The most important post-translational modification on this site is isoprenylation of conserved Cys residue. [205] RhoGDI is a down-regulator of Rho GTPases by preventing nucleotide exchange and controlling membrane association-dissociation cycling. It interacts with switch regions of Rho proteins and is blocking interactions to GEFs and GAPs by sterically blocking the binding sites. [204]

Flavin-binding photoreceptors regulate numerous biological functions in various organisms. They are categorized into three main groups: 1) LOV domains, 2) BLUF domains and 3) CRY domains.

Light-, oxygen- or voltage-sensitive (LOV) domains are photosensory units of phototropins, harboring a non-covalently bound flavin mononucleotide (FMN) chromophore. [79] Due to the FMN cofactor, upon blue-light excitation LOV domains undergo a photocycle via the excited triplet state of FMN. FMN excitation leads to a covalent bond formation between the C4(a) atom of FMN and a sulphur atom of a nearby conserved cystein, which represents the active signaling state of LOV domains. [93]

[1] Addition of hydrophobic molecules

A number of artificial photoswitches has been developed to extend the currently available optogenetic tools, and the LOV2 domain of phototropin1 from *Avena sativa* is often used in these artificial photoswitches. [206] One of these artificial photoswitches was a LOV2-Rac1 fusion construct, which was successfully used to control the motility of living cells. Photoactivatable Rac1 (PA-Rac1) was reversibly activated with localized blue light to precisely control generation of cell protrusions and ruffling. [98] The LOV2 domain of the construct serves as a photo-sensor and, upon activation, sterically releases the Rac1 domain, which is then able to interact with downstream effectors. [98] Upon activation, the signal is propagated through the LOV2 domain to the J$\alpha$ helix. This causes disruption of the J$\alpha$ helix. J$\alpha$-unfolding releases the steric inhibition of Rac1, leading to the activation of Rac1. [98]

In the present study, we elucidate the early stages of the signal transduction pathway of the Rac1 and Rac1b domains at the molecular level by molecular dynamics simulations. We also present the early signaling mechanisms of p67$^{phox}$-, RhoGDI- and LOV2-constructs with Rac1 and Rac1b domains, using the non-invasive molecular dynamics (MD) simulation method. We show similarities and differences in molecular mechanisms between Rac1 and Rac1b domains. We further solidify presented mechanisms and extend them to other Rac1/Rac1b complexes that regulate both physiological and pathological key processes. Details of the simulations, electrostatics calculations as well as of the generation of all of the starting structures are given in the Methods section.

## 6.2 Results

According to the study by Sondek *et al.* [173] there are 3 key regions responsible for effector binding control in Rac1: switch 1, switch 2 and p-loop. Hence, the analyses of our simulations will focus on these 3 important regions but will also include information on signal propagation in other parts of the protein. In addition, the analyses will also focus on guanosine phosphate as well as behavior of magnesium ion. When discussing Rac1b simulations we will focus on the same crucial regions, with additionally considering the 19-amino acid insertion. When analyzing LOV2-Rac1b simulations, we will also analyze the behavior of the calcium ion present in the structure.

All numberings used in residue identifications are readjusted original numberings of their respective structures to match the isolated Rac1 and Rac1b structural designations. These readjustments are done for easier comparison of the structures.

### 6.2.1 Rac1 and Rac1b simulations

Rac1 activation, by the introduction of GTP instead of GDP in the structure of Rac1, leads to a fast and direct signal propagation. The signal is transmitted to the 3 important regions: switch 1, switch 2 and p-loop.

Introduction of an additional phosphate group, through the nucleotide exchange, causes the p-loop to stay in place but it also causes it to form a hydrogen bond with the third phosphate group of the GTP. This does not happen, however, in the GDP simulation. The difference in p-loop binding is shown in figure 6.2.1a, specifically the binding between the GTP and Ala13 residue, where the Rac1-GDP (inactive) simulation is shown in black, while the Rac1-GTP (active) simulation is shown in red. Upon the GDP/GTP replacement, the additional phosphate group causes both the switch 1 and switch 2 to move inwards, towards the GTP, making a connection to the phosphate tail. This movement has an additional consequence, which is the hydrogen bond formation between switch 1 and switch 2 in the active state, which is not present in the inactive structure. Hydrogen bond formation between switch 1 and switch 2 is shown in figure 6.2.1b. The figure shows contacts formed between the residues Ala59 and Val36. In both states the magnesium ion is in an octahedral geometry, complexed with 2 water molecules, Thr17 and phosphate groups. In the inactive state, the magnesium ion is bound to both phosphate groups with 3 oxygens binding to the $Mg^{2+}$ ion, where one phosphate group binds through two oxygen atoms. In the active state, however, magnesium is complexed by all 3 phosphate groups, where each of the phosphate groups binds with $Mg^{2+}$ ion with a single oxygen atom. To accommodate for this change in complexation arrangement, the magnesium ion shifts slightly towards the third phosphate group, in the active state, thus making a more stable complex. The threonine residue that binds the magnesium ion lies on the $\alpha$1 helix, adjacent to the switch 1 region. The magnesium complex and its immediate surrounding in the active state are shown in figure 6.2.1d. Lastly, the replacement of GDP by GTP, makes the $\beta$2 and $\beta$3 sheets to move towards the $\alpha$5 helix. This is shown in figure 6.2.1c through the interactions of Asp47 with Asp170. There are no significant folding or unfolding changes induced by the GDP/GTP replacement.

In the Rac1b simulations we see mostly different behavior when compared to the Rac1 simulations. The signal is rapidly transmitted to all 4 of the crucial regions: p-loop, switches 1 and 2 and 19-amino acid insertion.

When comparing the Rac1b inactive and active simulations, the introduction of a third phosphate group induces several structural changes. The p-loop exhibits the same behavior like in the Rac1 simulations. It stays in place while forming a hydrogen bond with the GTP phospho-tail. The switch 1 region moves outwards, away from the GTP phosphate groups. Switch 2 forms a hydrogen bond with GTP tail. Due to this connection, the lower part of switch 2 stays locked in-place, while the upper part detaches from the rest of the protein. This is shown in figure 6.2.3a as a contact between the Asp10 and Gln60 residues. A hydrogen bond is established

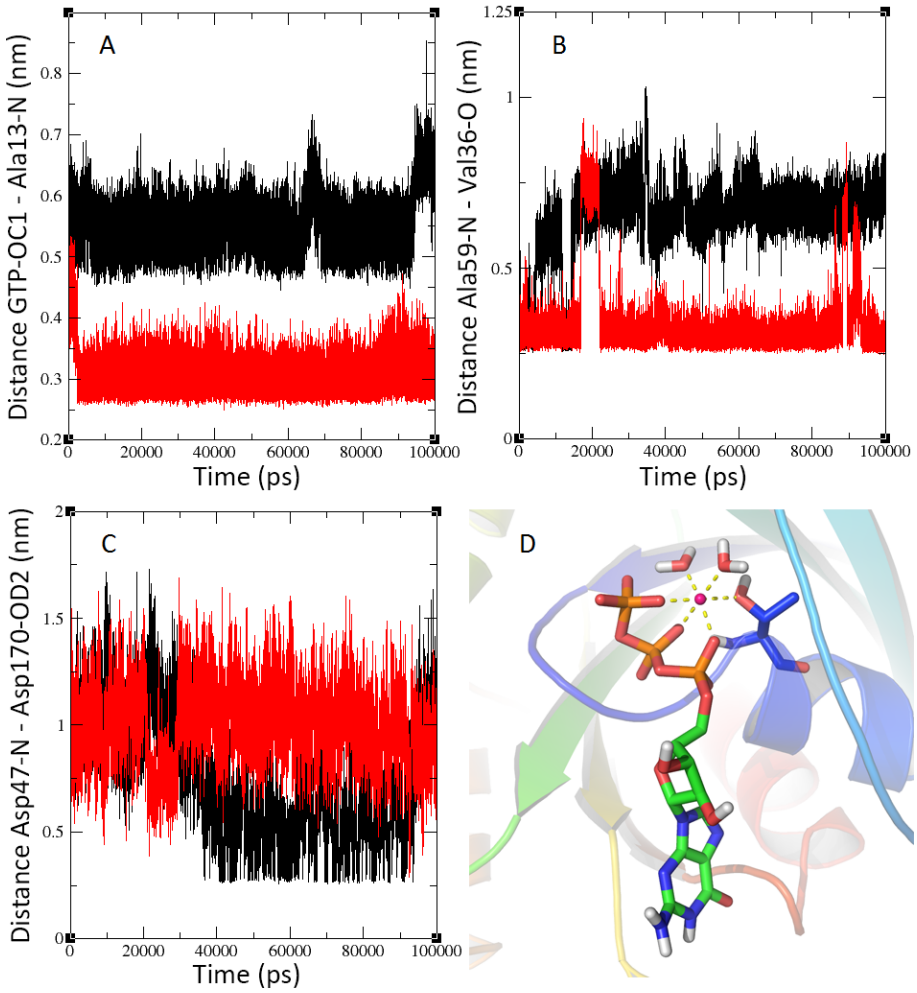


**Figure 6.2.1** Comparison of characteristic atom-pair distances (A) between the p-loop and GTP, (B) of hydrogen bond formation between switches 1 and 2, (C) of the movement of the $\beta$2 and $\beta$3 sheets towards the $\alpha$5 helix. Panel (D) shows the active state structure of magnesium complex and its surrounding after 100 ns of simulation. The Rac1 inactive states are shown in black, while the Rac1 active states are shown in red in all figures.

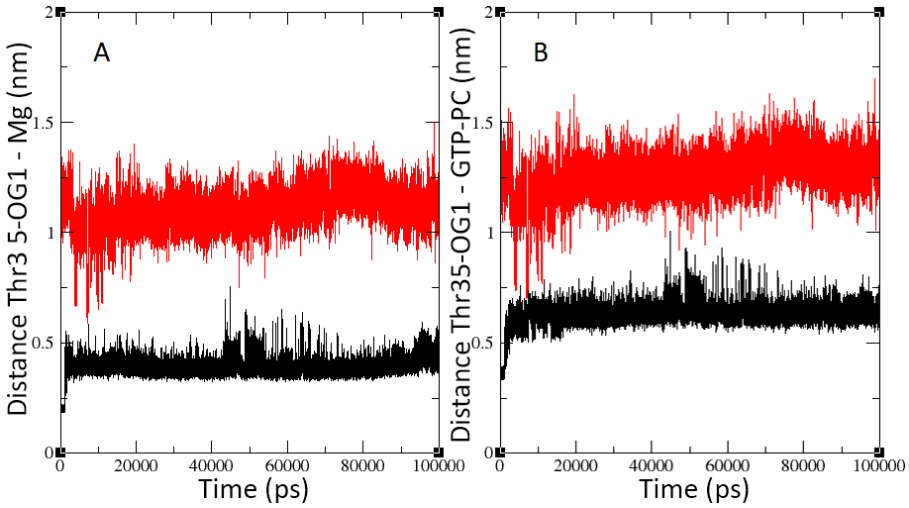


**Figure 6.2.2** Comparison of characteristic atom-pair distances of (A) Thr35 and the magnesium ion and (B) Thr35 and the $\gamma$-phosphate group of GTP. The Rac1 active state is shown in black, while the Rac1b active state is shown in red in all figures.

between switches 1 and 2 in the active state, while no such connection is formed in the inactive state. The switch 1 - switch 2 hydrogen bond is shown in figure 6.2.3b through the bonding between residues Thr35 and Gln60. The Thr35 residue and its connection to the magnesium ion is shown in figure 6.2.2a, while 6.2.2b shows the connection between Thr35 and $\gamma$-phosphate from GTP, in the Rac1 and Rac1b active states. Interestingly, the 19-amino acid insertion loop shows significant detachment from the protein into the solvent. This occurs however, only in the active state structures. The detachment of the 19-amino acid insertion is shown in figure 6.2.3c where the interaction of Asn38 with Asp82 is shown as a function of simulation time. The magnesium complex in the inactive state shows the same arrangement and same complexation partners as in Rac1-GDP simulation. It is bound to two phosphate groups, Thr17 and 2 water molecules. However, in the active state there is a significant change. The $Mg^{2+}$ ion is bound to two terminal phosphate groups, a single water molecule and both oxygens of Asp56, keeping its octahedral geometry. The residue that binds to the $Mg^{2+}$ ion changes from Thr17 on the $\alpha$1 helix to Asp56 that is on the $\beta$3 sheet. This change in complexation partners can explain the movement of $\alpha$1, $\beta$2 and $\beta$3 regions. Contrary to the Rac1 simulations, here the $\beta$2 and $\beta$3 sheets move outwards, while the $\alpha$1 helix shifts away from the protein. The $Mg^{2+}$ complex and its surroundings are shown in figure 6.2.3d.

Interesting to note here is that the movement of switch1, switch 2, $\beta$2 and $\beta$3 sheets in the Rac1b active simulations is opposite to the movement of the same regions in the Rac1 simulations. In Rac1b, these regions are going outwards, towards the solvent, while in Rac1 they are moving inwards, towards the GTP and away from the water molecules.

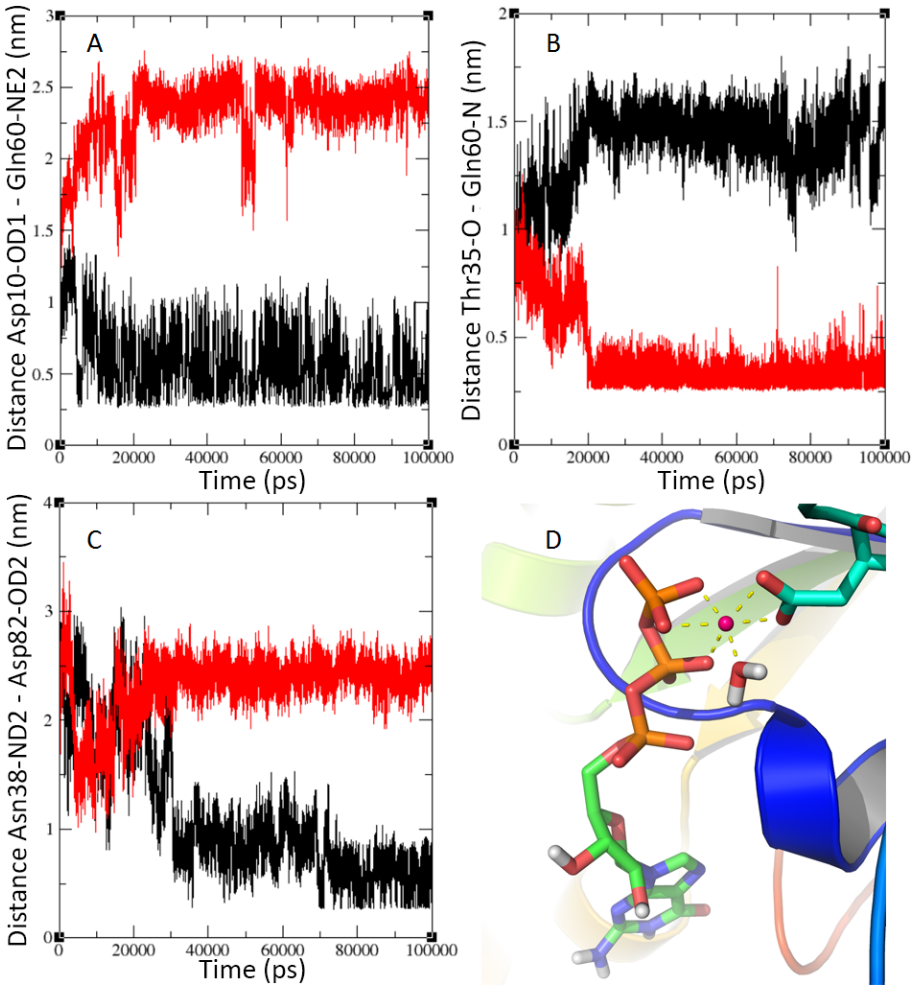


**Figure 6.2.3** Comparison of characteristic atom-pair distances of (A) switch 2 detachment from the protein, (B) hydrogen bond formation between switches 1 and 2, (C) the movement of the 19-amino acid insertion loop towards the solvent molecules. (D) shows the active state structure of the magnesium complex after 100 ns. The Rac1b inactive states are shown in black, while the Rac1b active states are shown in red in all figures.

### 6.2.2 RhoGDI-Rac1 simulations

Performing a direct GDP/GTP exchange in the RhoGDI-Rac1 complex induces a slow signal transduction localized only in the Rac1 domain. The signal is transmitted to the important regions of Rac1. The behavior of the Rac1 domain within the RhoGDI-Rac1 complex is comparable to the behavior exhibited by the isolated Rac1 domains.

Direct GDP by GTP nucleotide exchange causes the p-loop to form a hydrogen bond with GTP while staying in place. There is no hydrogen bond formed between the p-loop and GDP. This difference in hydrogen bonding, between the inactive and active states of the complex, is shown in figure 6.2.4a as a contact between nucleotide and Ala13$_{Rac1}$. Introduction of the third phosphate group causes the switch 1 to form a side-connection with the rest of the protein which causes the movement of the switch 1 region inwards, towards the GTP. This connection, specifically between Leu160$_{Rac1}$ and Pro29$_{Rac1}$ is shown in figure 6.2.4b. Due to the binding orientation of the RhoGDI domain, the switch 2 region of Rac1 is locked in an outwards position, away from the GTP, similar to the position of the switch 2 region in Rac1b. This locked conformation is the reason for the lack of hydrogen binding between switch 1 and switch 2 in both active and inactive states. The center of mass distances of RhoGDI and Rac1 domains as a function of simulation time are shown in figure 6.2.4c. Magnesium, in the inactive state, forms a complex with Asp56$_{Rac1}$, a single phosphate group from GDP and 2 water molecules. Different to this, in the active state, $Mg^{2+}$, while keeping octahedral geometry, is complexed by Thr17$_{Rac1}$, two phosphate groups from GTP and three water molecules. The change in complexation partners is caused by the locked orientation of the switch 2 region as well as the movement of the other regions of Rac1, upon activation. The geometry of magnesium complex, in the active state, is shown in figure 6.2.4d. The GDP/GTP exchange also induces movement of the $\beta$2 and $\beta$3 sheets inwards, toward the $\alpha$5 helix.

The signal was not propagated further from Rac1 within the simulation time of 100 ns. The connection between Thr35$_{Rac1}$ and Asp215$_{GDI}$, responsible for the inhibition by the GDP/GTP exchange [207], is stable over the simulation time, independent of the bound nucleotide. The interaction surface between RhoGDI and Rac1 also does not change throughout the simulations.

### 6.2.3 p67$^{phox}$-Rac1 and p67$^{phox}$-Rac1b simulations

Upon GDP/GTP exchange in p67$^{phox}$-Rac1 simulations, the p-loop and switch 1 show no significant change. Switch 2 moves towards GTP and shows partial unfolding in the region of residues 62-67. These changes do not occur in the GDP simulation. Similar to the Rac1 simulations, the $\beta$2 and $\beta$3 sheets move towards the $\alpha$5 helix, in the active state. The magnesium ion in the inactive state is complexed

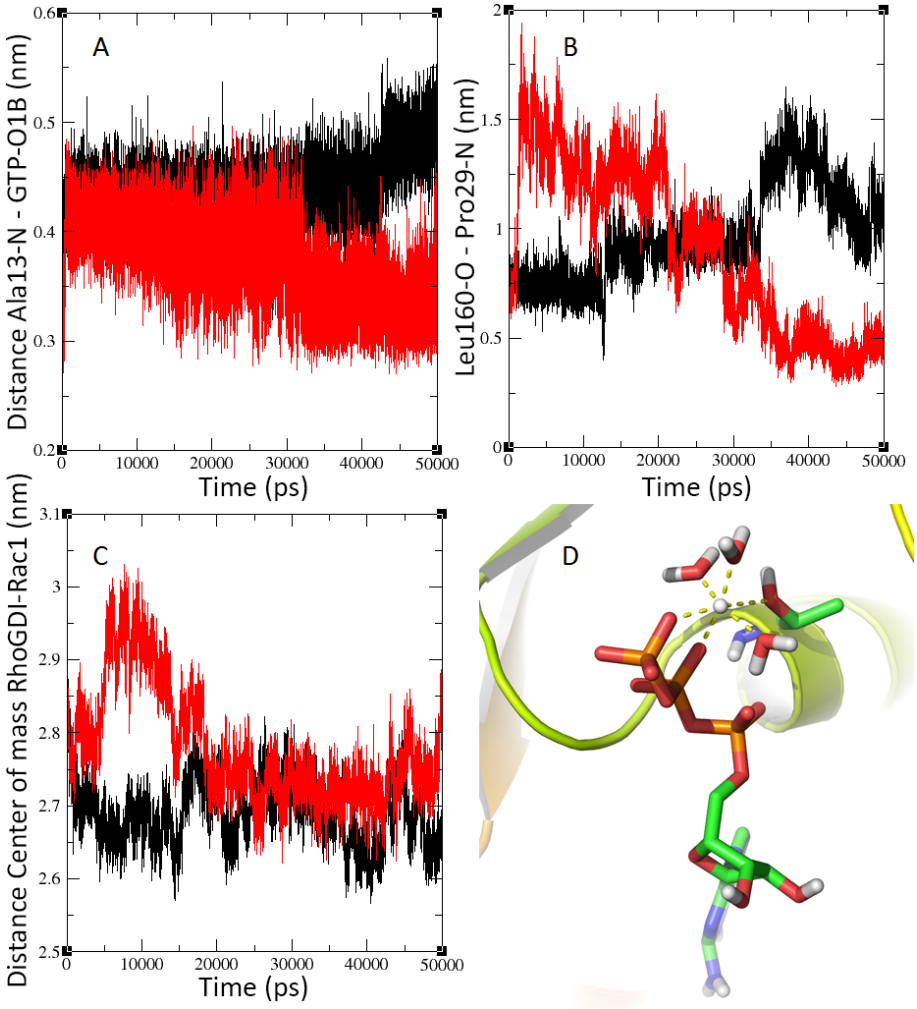


**Figure 6.2.4** Comparison of characteristic atom-pair distances (A) between the p-loop and the nucleotide, (B) of tightening of the connection between switch 1 and the rest of the protein, (C) of the center of mass movement of RhoGDI and Rac1 domains and (D) shows the active state structure of the magnesium complex. The inactive states are shown in black, while the active states are shown in red in all figures.

with both phosphate groups, Thr17$_{Rac1}$, Thr35$_{Rac1}$ and one water molecule. The complexation partners of $Mg^{2+}$ do not change upon activation of the Rac1 domain in the fusion protein.

The signal was propagated to the interaction surface between p67$^{phox}$ and Rac1. Figure 6.2.5a shows the change in the distance between the two subunits. There is a separation of two domains, over the simulation time, in the active state. Contrary to this, the inactive state shows stronger binding between the units, where the center of masses come closer together.

Arg102$_{p67}$ is a key residue involved in the interaction surface. It interacts with Ala159$_{Rac1}$, Leu160$_{Rac1}$, Asn26$_{Rac1}$ and Ser22$_{Rac1}$. Additionally, Asn104$_{p67}$, Asp108$_{p67}$ and Asp67$_{p67}$ are also involved in the binding between the two domains, with their counterparts from the Rac1 domain, Gln162$_{Rac1}$, Thr25$_{Rac1}$ and Gly30$_{Rac1}$. The interaction of Arg102$_{p67}$ with Asn26$_{Rac1}$ and Ser22$_{Rac1}$ is shown in figures 6.2.6a and 6.2.6c, respectively. Figure 6.2.6e shows the interaction between Asp67$_{p67}$ and Gly30$_{Rac1}$.

Upon activation, p67$^{phox}$-Rac1b simulations show a different behavior compared to the p67$^{phox}$-Rac1 simulations. Both p67$^{phox}$-Rac1b simulations show results that are similar to those of the isolated Rac1b. While the p-loop and the switch 1 regions do not change, the switch 2 region moves outwards, away from GTP, while it does not unfold. There is no change in the $\beta$2 and $\beta$3 sheets. The 19-amino acid insertion contracts and makes tighter connection with the rest of the protein. It also makes a hydrogen bond with $\beta$3 in the active p67$^{phox}$-Rac1b instead of $\beta$2 in the inactive p67$^{phox}$-Rac1b. In the inactive state, magnesium forms a complex with two phosphate groups, Thr17$_{Rac1}$, Thr35$_{Rac1}$ and one water molecule. Addition of the phosphate group causes a change in the binding partners. In the active state it binds to two phosphate groups, Asp56$_{Rac1}$ and a single water molecule.

The center of mass distances, shown in figure 6.2.5b, show no significant change in the unit distances when comparing the p67$^{phox}$-Rac1b inactive and active states. The residues Ser22$_{Rac1}$, Thr25$_{Rac1}$, Asn26$_{Rac1}$, Gly30$_{Rac1}$ and Glu31$_{Rac1}$ are the Rac1b residues involved in the interaction between Rac1b and p67$^{phox}$. Interaction of Arg102$_{p67}$ with Ser22$_{Rac1}$ and Asn26$_{Rac1}$ show no large changes when comparing the two states. The two interactions are shown in figures 6.2.6b and 6.2.6d. The connection between Asp67$_{p67}$ and Gly30$_{Rac1}$ stretches, when going to the active state. This connection is shown in figure 6.2.6f. The Glu31$_{Rac1}$ residue changes orientation and gets more buried into the p67$^{phox}$ domain. Residues Ser37$_{p67}$, Asp67$_{p67}$, Arg102$_{p67}$, Asn104$_{p67}$ and Asp108$_{p67}$ are crucial residues involved in the p67$^{phox}$-Rac1b interaction surface. Protein-protein interactions between p67$^{phox}$ and Rac1b domains and their subsequent changes upon activation of Rac1b, are shown in figure 6.2.7. Specifically the interactions between the Gly81$_{Rac1}$ and Thr24$_{Rac1}$ residues are shown in figure 6.2.7a, while figure 6.2.7b shows the interactions between residues Lys82$_{Rac1}$ and Tyr23$_{Rac1}$. Contacts formed between Asn104$_{p67}$ and Gln181$_{Rac1}$ residues can be

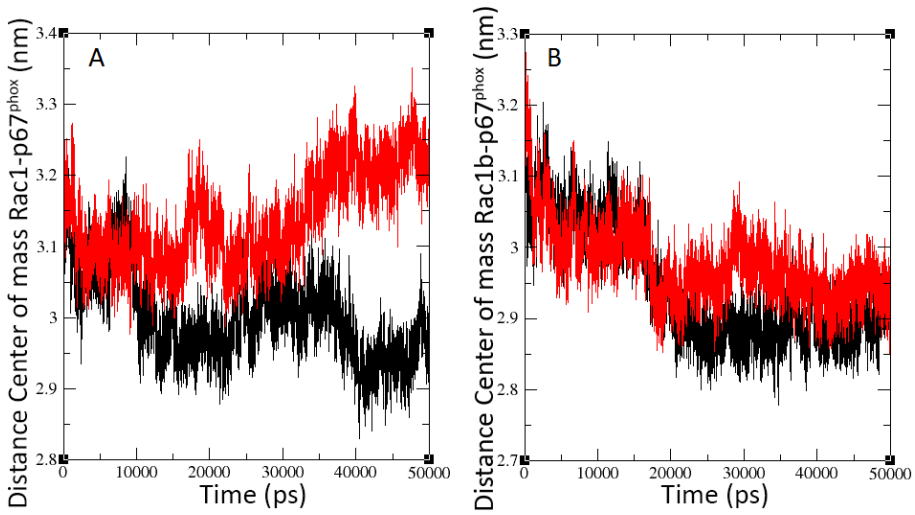


**Figure 6.2.5** Comparison of center of mass distances between (A) Rac1 and p67$^{phox}$ domains and (B) Rac1b and p67$^{phox}$ domains. The inactive states are shown in black, while the active states are shown in red in all figures.

seen in figure 6.2.7c. The contacts of Ala27$_{Rac1}$ with Gln181$_{Rac1}$ residue are shown in figure 6.2.7d.

### 6.2.4 LOV2-Rac1b simulations

We performed four different simulations of LOV2-Rac1b system: combinations of light and dark states of the LOV2 domain with active and inactive states of the Rac1b domain. When comparing the dark-inactive to the dark-active state (only activating the Rac1b domain), there are several changes induced in the Rac1b domain. These changes do not extend to the LOV2 domain. In the Rac1b domain, the p-loop stays in place, while both switch 1 and switch 2 move outwards, simultaneously making a hydrogen bond between them. This hydrogen bond between switch 1 and switch 2, in both active and inactive states, is shown in figure 6.2.8a as an interaction between the Asp56$_{Rac1}$ and Phe36$_{Rac1}$ residues. The 19-amino acid insertion region extends further, looping around the LOV2 domain so that the LOV2 domain sits tightly on top of the 19-amino acids insertion. The $\beta$2 and $\beta$3 sheets move outwards as well, breaking their connection to the $\alpha$5 helix. Interestingly, the complexation partners of the $Mg^{2+}$ ion change significantly. In the dark-inactive state, magnesium is complexed to 2 phosphate groups, Thr17$_{Rac1}$ and 2 water molecules from the solvent. In the dark-active state, however, magnesium complexes two phosphate groups, Thr35$_{Rac1}$ and Asp56$_{Rac1}$ with no water molecules bound.

There are no significant changes induced in the LOV2 domain of the LOV2-Rac1b fusion protein, upon activating only the Rac1b domain. The J$\alpha$ helix as well as the FMN surrounding show no major movement or events happening during the

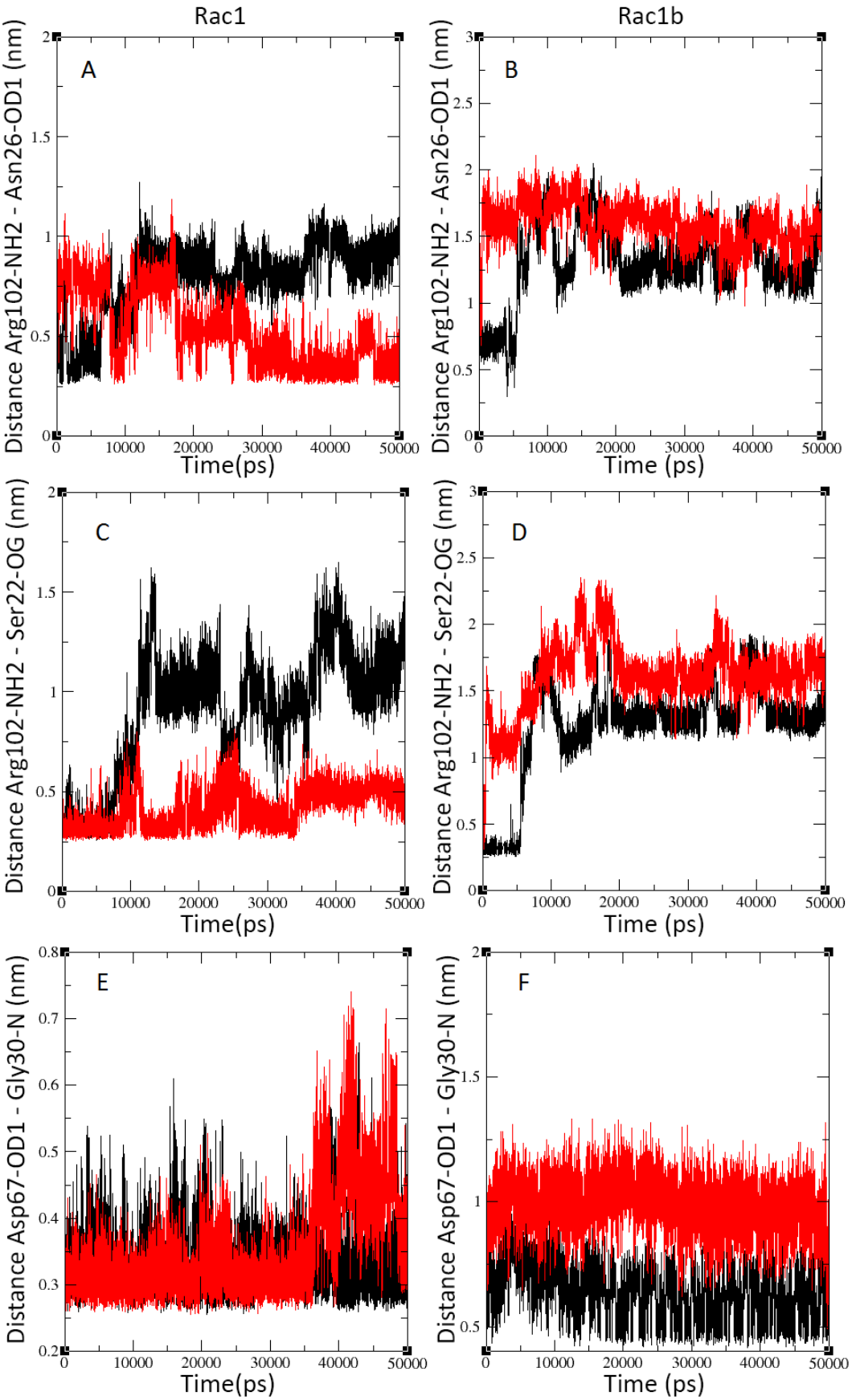


**Figure 6.2.6** Comparison of characteristic atom-pair distances between residues (A) $Arg102_{p67}$ and $Asn26_{Rac1}$ in $p67^{phox}$-Rac1, (B) $Arg102_{p67}$ and $Asn26_{Rac1}$ in $p67^{phox}$-Rac1b, (C) $Arg102_{p67}$ and $Ser22_{Rac1}$ in $p67^{phox}$-Rac1, (D) $Arg102_{p67}$ and $Ser22_{Rac1}$ in $p67^{phox}$-Rac1b, (E) $Asp67_{p67}$ and $Gly30_{Rac1}$ in $p67^{phox}$-Rac1 and (F) $Asp67_{p67}$ and $Gly30_{Rac1}$ in $p67^{phox}$-Rac1b. The inactive states are shown in black, while the active states are shown in red in all figures.

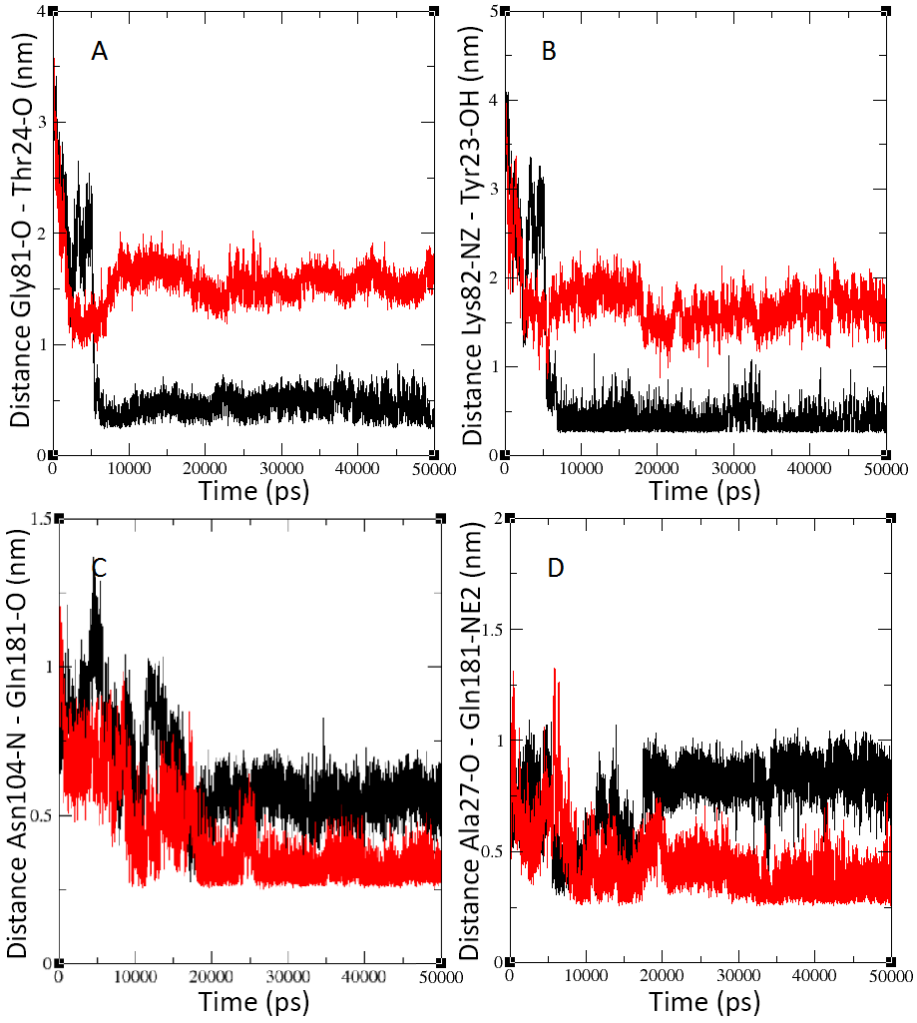


**Figure 6.2.7** Comparison of characteristic atom-pair distances between (A) Gly81$_{Rac1}$ and Thr24$_{Rac1}$, (B) Lys82$_{Rac1}$ and Tyr23$_{Rac1}$, (C) Asn104$_{p67}$ and Gln181$_{Rac1}$ and (D) Ala27$_{Rac1}$ and Gln181$_{Rac1}$ residues. The inactive states are shown in black, while the active states are shown in red in all figures.

simulations. However, there is a big change in the binding of the calcium ion, which gets released into the solution and does not bind to the protein anymore.

Activation of the LOV2 domain, by forming a covalent bond between FMN and the adjacent cystein residue, induces characteristic changes in the LOV2 domain. The complete signaling pathway of LOV2 [16] seems to be conserved in the LOV2-Rac1b fusion protein. The most prominent changes induced, going from dark-inactive to light-inactive structures, are breakage and unfolding of the J$\alpha$ helix as well as unfolding of the A'$\alpha$ helix. There are significant torsions of the G$\beta$, H$\beta$ and I$\beta$ sheets together with the unfolding of the I$\beta$ sheet. Comparison of the structures as well as unfolding and breakage of J$\alpha$ upon activation is shown in figure 6.2.8b. Another important feature of the light-induced molecular mechanism of LOV2 is the rotational reorientation of the glutamine residue. This difference in states is visualized in figures 6.2.8c and 6.2.8d. It has been shown numerous times [16, 57–61] that this flip of the $Gln513_{LOV2}$ residue is a critical step in LOV2 signaling.

Additionally, activating the Rac1b domain in the light-activated LOV2 domain, causes switch 1 to move outwards and switch 2 to move inwards, thus making a hydrogen bond between them. The p-loop, the 19-amino acid insertion and both $\beta$2 and $\beta$3 regions show no significant change when comparing light-inactive versus light-active simulations.

In the same fashion, activating the LOV2 domain connected to an activated Rac1b domain, causes a number of characteristic changes on the contact interface between the LOV2 and Rac1b domains. Most notably, when comparing dark-active to light-active states, the breakage of the J$\alpha$ helix causes a significant contraction of the 19-amino acid insertion region. There is also noticeable twisting of the $\beta$2 and $\beta$3 regions, which causes a movement of switches 1 and 2 towards the solvent, making them more exposed.

The center of mass distances of the Rac1b and LOV2 subunits of the protein construct are shown in figure 6.2.9. As can be seen in the graphs, the distance between the two units is dictated by the Rac1b activation state. This is contrary to the expected LOV2-control of the distance between the units. In the cases of dark-inactive and light-inactive states, the distances stay constant over the simulation times, while in the dark-active and light-active states, the distance between the LOV2 and Rac1b subunits decreases over time.

Figure 6.2.10 shows electrostatic potentials of the LOV2 and Rac1b subunits in different simulations. The most noticeable regions are the 19-amino acids insertion, from the Rac1b unit, which is positively charged, while the saddle area of the LOV2 domain, that sits on top of the 19-amino acid insertion, is negatively charged. These oppositely charged regions further strengthen the interaction between the subunits, through electrostatic contacts.

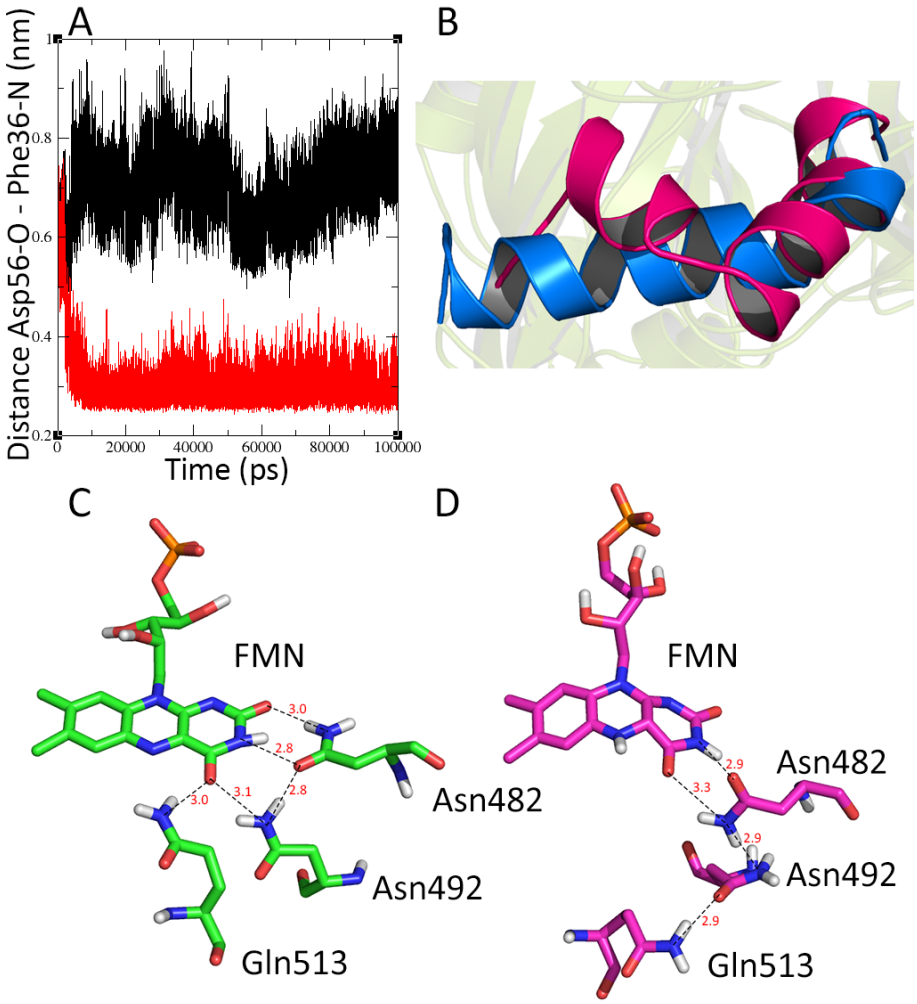


**Figure 6.2.8** Structural changes of LOV2-Rac1b fusion protein regions. (A) Hydrogen bond distance between switch 1 and switch 2 regions of the Rac1b domain. The inactive state is shown in black, while the active state is shown in red. (B) Structural changes of the J$\alpha$ helix in the dark (blue) and light (red) states. Rest of the LOV2-Rac1b protein shown in green in the background. (C) Hydrogen bonding network around the FMN in the dark state, including the Gln513$_{LOV2}$ structural orientation together with Asn482$_{LOV2}$ and Asn492$_{LOV2}$. Structural changes of the same residues in the surrounding of the FMN upon activation of the LOV2 domain are shown in (D).

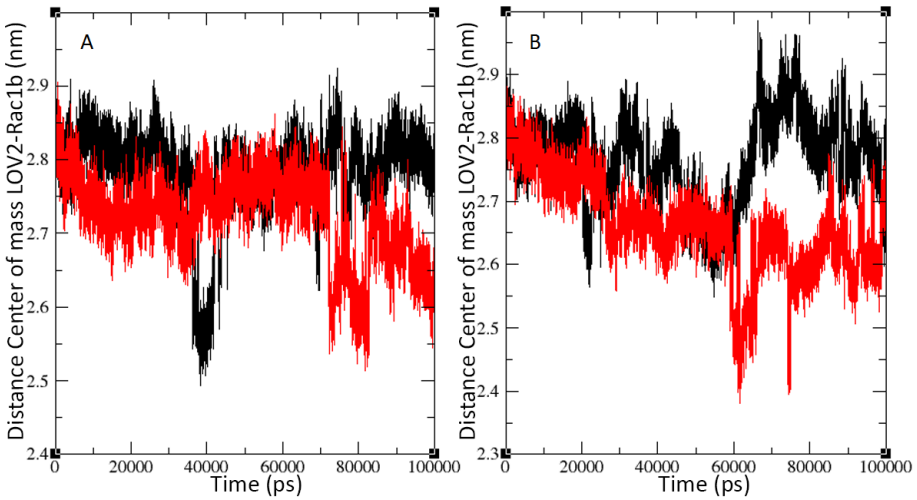


**Figure 6.2.9** Comparison of the center of mass distances between Rac1b and LOV2 domains in (A) dark-inactive (black) and dark-active states (red) and (B) light-inactive (black) and light-active (red) states.

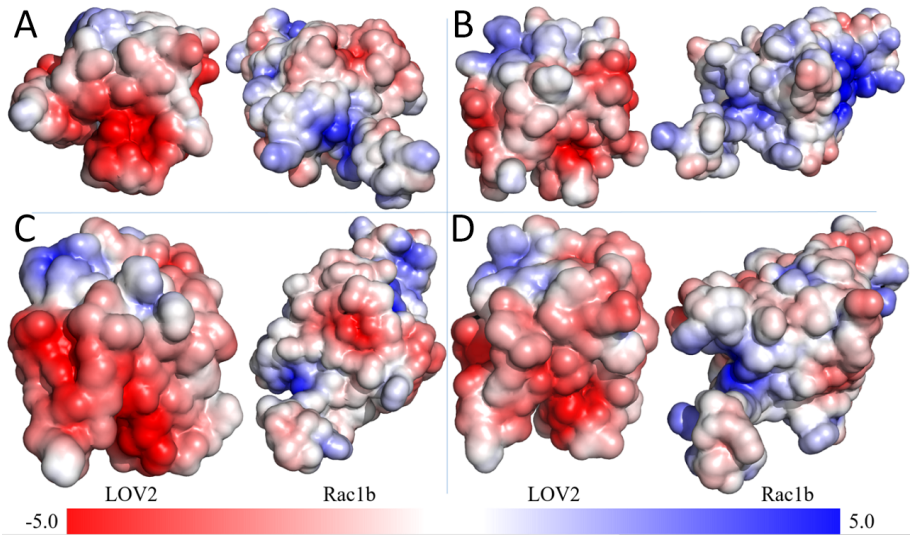


**Figure 6.2.10** APBS surfaces of the LOV2 and the Rac1b units of the 100ns structures from (A) dark-inactive, (B) dark-active, (C) light-inactive and (D) light-active simulations.

## 6.3 Discussion

To analyze the early signaling mechanism of Rac1 and its spliced variant, Rac1b, we performed a number of simulations of these two proteins and their protein complexes and fusion constructs with other domains, in their inactive and active states. We will start the discussion with the analysis of the molecular mechanisms of isolated Rac1 and Rac1b domains, which we will further expand on with the analysis of their constructs.

### 6.3.1 Rac1 and Rac1b signaling mechanism

While the structures of Rac1 and Rac1b differ by only the 19-amino acid insertion at the end of the Switch 2 region, their early signaling mechanisms are distinct and quite divergent. When aligned, the structures of Rac1 and Rac1b show a high degree of similarity (total RMS of 0.371 Å).

There are 3 important regions of Rac1 reported in the literature [173]: p-loop, switch 1 and switch 2. Rac1b additionally has a 19-amino acid insertion region that is crucial for its signal propagation. All of these three regions (four in the case of Rac1b) are located around guanosine nucleotide. So the mechanism induced by the exchange of GDP with GTP is fast and the signal is directly propagated to these important regions.

The p-loop plays a crucial role in phosphate binding and coordination of the magnesium ion, through the side chain of Thr17. In Rac1 simulations the p-loop shows no significant structural change. It does form a hydrogen bond with the guanosine base in the active state, however, the reason for this is the movement of the base rather than the reorientation of the residues of the p-loop. The p-loop of Rac1b shows the same behavior as the p-loop of Rac1. It shows no significant change upon activation except that the active state forms a hydrogen bond with GTP.

The switch 1 and switch 2 regions are crucial regions controlling downstream mechanism through partner binding. In the Rac1 simulations, they both show inward movement, towards GTP, while simultaneously forming a hydrogen bond between them (closed conformation). As these two regions are responsible for binding of Rac1 to its reaction partners, the tighter binding of these two regions to the protein core suggests decreased binding of Rac1 to its binding partners. Interestingly, Rac1b shows opposite behavior. Switches 1 and 2 move outwards, into the solution (open conformation), increasing the contact surface with the solvent and making them more flexible. This weaker binding to the protein core, going from a closed conformation in Rac1 to an open conformation of switches 1 and 2 in Rac1b, points towards a more reactive nature of Rac1b and easier binding to its binding partners.

According to Fiegen and coworkers [175], the catalytic residue Gln60 in switch 2 is an important residue for impaired the GTPase activity of Rac1b. In our simulations

we see that the Gln60 residue from the switch 2 region plays an important role. We show in figure 6.2.3b that the contact that Gln60 makes with Thr35 (switch 1 region) further triggers the movement of switch 1 region. The high flexibility of the switch 2 region induces a movement in the switch 1 region. This high flexibility of switch 2 is in turn induced by the introduction of the 19-amino acid insertion in the structure of Rac1b.

In the Rac1b inactive states, we see no interaction between switch 1 and switch 2 in our simulations. In the active state however, switch 1 and switch 2 form a hydrogen bond between them, locking the lower regions of the two switches while the upper parts form open conformations increasing their affinity for partner binding.

In their work [175], Fiegen *et al.* point out that the Phe37 residue of switch 1 lies in a hydrophobic cleft, formed by the side chains of Thr58, Tyr64, Leu67 and Arg68 from switch 2. However, in our simulations there is a hydrophobic pocket formed between the side chains of Phe37, Leu67 and Tyr64, while the Arg68 and Thr58 side chains are rotated outwards from the pocket and are not apart of this hydrophobic cleft. The Val36 side chain makes hydrophobic contacts with Tyr64 and Leu67 residues. While the high mobility of the switch 2 and 19-amino acid insertion regions break these hydrophobic contacts, there is a hydrogen bond formed between switches 1 and 2, stabilizing the open conformation of switch 1.

The octahedral geometry of the magnesium ion is highly conserved throughout small GTPases [175, 208–210] and it shows the same octahedral geometry in all of our simulations. The magnesium ion is complexed with the $\beta$-phosphate group of the nucleotide, the Thr17 side chain and water molecules in the inactive state. This does not change upon activation of Rac1, by the nucleotide exchange. Contrary to this, the magnesium ion in Rac1b shows a change upon activation, where the side chain of the Thr17 residue is replaced with the side chain of the Asp56 residue, while keeping the contacts with the solvent and the guanosine base. This change in complexation of the magnesium ion in inactive and active states has been described before [208–210] for other members of Ras protein family.

The region only present in Rac1b, the 19-amino acid insertion, shows a clear extension upon GDP/GTP exchange. This behavior of the 19-amino acid insertion is present in all of the simulations of Rac1b constructs, however it is controlled by the binding position and orientation of the binding partner. The 19-amino acid insertion induces higher flexibility in the switch 2 region.

The open conformations of the switch 1 and switch 2 regions have an additional consequence. In Rac1 simulations, this movement causes the $\beta$2 and $\beta$3 sheets to move inwards and form a connection to the $\alpha$5 helix. In simulations of Rac1b however, this behavior is reversed, showing a detachment of these two beta sheets. Also, the $\alpha$1 helix exhibits a detachment from the p-loop in Rac1b, while this does not occur in Rac1 simulations.

The open conformation of switch 1 additionally causes the contacts between Thr35 and the magnesium ion, Thr35 and the $\gamma$-phosphate group of GTP to be disrupted

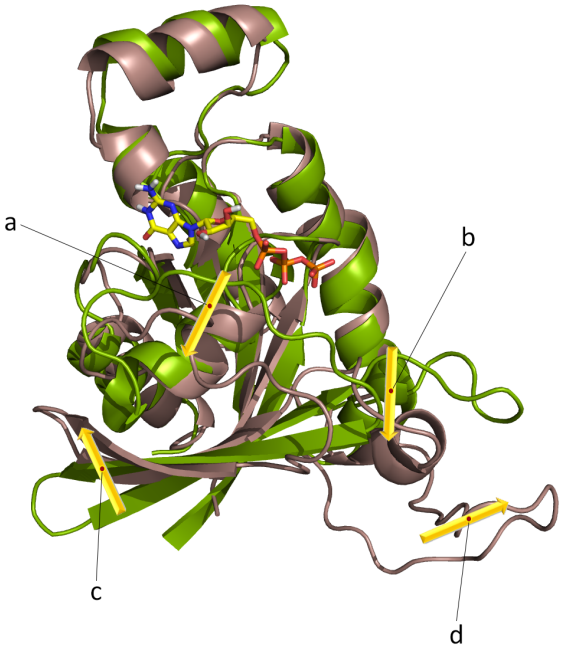


**Figure 6.3.1** Superposition of the Rac1 and Rac1b 100 ns active state structures. The Rac1 structure is colored green, while the Rac1b structure is colored dark-red. GTP is shown as sticks and the structural changes are shown with yellow arrows pointing in the direction of the change going from the Rac1 to the Rac1b structures. Changes introduced a) opening of the switch 1 region, b) opening of the switch 2 region, c) changes of the $\beta$-hairpin, and d) movement of the 19-amino acid insertion into the solution.

in the active state of Rac1b, when compared to Rac1, as shown in figure 6.2.2a and 6.2.2b. This in turn causes a drastic reduction in nucleotide affinity by allowing the release of nucleotide from the protein core. This structural change between Rac1 and Rac1b can provide an explanation for the rapid nucleotide dissociation of Rac1b. [175] In their study, Orlichenko and colleagues [168] found that GIT1, RhoGDI and IQGAP effectors have higher affinity towards Rac1 as compared to Rac1b. The reason for this could be the open conformation of switch 1 and switch 2 in Rac1b compared to Rac1. Kumar *et al.* [211] show that the gain-of-function F28L point mutation in Rac1 destabilizes the switch 1 region which they propose is the crucial reason for the loss of GTPase activity of the protein. This destabilization can be further extended and applied to the open conformation of switch 1 pointing to the same reason (the destabilization of switch 1) being the cause of the loss of GTPase activity in Rac1b.

The main structural changes described above are illustrated in figure 6.3.1, where the 100 ns active-state structures of Rac1 and Rac1b are superimposed. Movements of the regions discussed above are shown by arrows in the direction of the changes, going from Rac1 to Rac1b structures. An exception to this is the movement of the 19-amino acid insertion region in Rac1b as it is not present in the Rac1 structure. The movement of the 19-amino acid insertion is drawn by comparing the 100 ns structures of Rac1b in the inactive and active states.

### 6.3.2 RhoGDI-Rac1 signaling mechanism

It has been shown [195] that Rac1b does not bind to RhoGDI and PAK1. Hence, we performed RhoGDI-Rac1-GDP and RhoGDI-Rac1-GTP and no Rac1b construct simulations in an effort to further show the signaling pathway of Rac1 that originates from the GDP/GTP exchange.

Nucleotide exchange induces signals localized to only the Rac1 domain. The signal propagation is much slower due to the interactions with the bound partner, which induced sterical tensions and slowed down the movement of different regions upon Rac1 activation. The behavior of Rac1 within the RhoGDI-Rac1 complex is comparable to the behavior of the isolated Rac1 domain. The p-loop shows no change while making a hydrogen bond to the phosphate tail of the GTP, while there is no hydrogen bond formed in the simulation of inactive systems. Switch 1 typically shows a closed conformation, while, due to the binding position and orientation of RhoGDI on Rac1, switch 2 is locked in an open conformation but with lower flexibility. This locked conformation is the reason for the lack of hydrogen bond contacts between switch 1 and switch 2.

The magnesium ion forms a complex with the side chain of $Asp56_{Rac1}$ in the inactive state, while in the active state this residue is replaced with $Thr17_{Rac1}$. This change in complexation partners, while keeping the octahedral geometry, is caused by the locked open conformation of switch 2 as well as the movement of switch 1. Due to the switch 2 being pulled out, because of the interaction with RhoGDI, there is more space for the water to penetrate inside the protein, so the magnesium ion forms a complex with 3 water molecules.

The magnesium complexation partners change over simulation time and are different to the reported [207] $Thr35_{Rac1}$ residue. All starting structures (generated using crystal structures as a template) contain $Mg^{2+}$ complexed with $Thr35_{Rac1}$. However, this complexation residue changes in our simulations. The reason for this change is a flip of $Thr17_{Rac1}$ that occurs in both active and inactive simulations, induced by the binding of the RhoGDI to the Rac1. Due to this structural change, the magnesium surrounding changes as well and the complexation role is taken up by the $Asp56_{Rac1}$ in the inactive state, and further replaced by $Thr17_{Rac1}$ residue in the active state. The same $Thr17_{Rac1}$ residue is also the complexation partner in the Rac1-GDP and Rac1-GTP simulations. This leads us to conclude that the sterical tension is released during the simulation, which allows the $Mg^{2+}$ ion to take the most optimal orientation within the protein structure, resembling the geometry in the Rac1 isolated domain.

As Hakoshima and coworkers reported [205], GDI makes contact primarily with switch 2 but also with the C-terminal region of switch 1. The connection between $Thr35_{Rac1}$ and $Asp215_{Rac1}$ is responsible for GDP/GTP exchange inhibition [207] but it is also responsible for magnesium coordination and consequently enhances the affinity towards the nucleotide [205]. This connection is stable over the simulation

time, irrelevant of which nucleotide is bound to the protein. The extensive contacts to the switch 2 region induce stabilization of the switch 1 region which leads to the inhibition of both [205] GDP dissociation and GTP hydrolysis.

Comparing all of these structural features of the RhoGDI-Rac1 complex with the features of Rac1 and Rac1b isolated domains, we can conclude that the behavior exhibited by the Rac1 domain is similar to that of the isolated Rac1 domain with the difference of switch 2 taking up an open conformation. The reason for this difference is the binding position and contacts that RhoGDI makes with the Rac1 domain, locking switch 2 in an open conformation. No significant change is observed in the distances of center of masses of the two domains all leads us to propose that in addition to the GDP/GTP exchange, there are other requirements such as protein binding, [212, 213] for the RhoGDI-Rac1 dissociation that are not represented in our simulations.

### 6.3.3 p67$^{phox}$-Rac1 and p67$^{phox}$-Rac1b signaling mechanism

The Rac1 and Rac1b bind to the N-terminal region of p67$^{phox}$. The binding interface between the Rac1 and Rac1b units with the p67$^{phox}$ unit contains four tetratricopeptide repeats (TPR). [201] The TPR region does not bind to the switch 2 region of Rac1 nor Rac1b. However, the protein-protein interface includes the regions of switch 1 and the loop between the $\beta$5 sheet and the $\alpha$5 helix.

Upon Rac1 activation of the p67$^{phox}$-Rac1 complex, the p-loop and switch 1 show no significant change. The most likely reason is that switch 1 and the loop between $\beta$5 and $\alpha$5 are used as an interaction surface between the two domains. Contrary to this but in accordance to already described mechanism of isolated Rac1, switch 2 shows a significant change by taking a more closed conformation. However, it also shows unfolding in the regions of the residues 62-67. These changes are not exhibited by the simulations of inactive state. As in the isolated Rac1 mechanism, the $\beta$2 and $\beta$3 sheets make a connection with the $\alpha$5 helix, in the active state. Since the protein-protein interaction surface is switch 1 and the $\beta$5-$\alpha$5 loop, the magnesium ion is complexed with both Thr17$_{Rac1}$ and Thr35$_{Rac1}$ in addition to the phosphate groups and a single water molecule. Upon activation of the Rac1 domain, the switch 1 does not change structurally and the complexation partners of the magnesium ion do not change as well.

The signal of Rac1 activation was propagated to the protein-protein interface during the time of the simulation. Asp102$_{p67}$ is a crucial residue in the interaction surface. It interacts with Ala159$_{Rac1}$, Leu160$_{Rac1}$, Asn26$_{Rac1}$ and Ser22$_{Rac1}$. Additionally, Asn104$_{p67}$, Asp108$_{p67}$ and Asp67$_{p67}$ are also involved in the binding between the two domains, with their counterparts from the Rac1 domain, Gln162$_{Rac1}$, Thr25$_{Rac1}$ and Gly30$_{Rac1}$, respectively.

Activation of the Rac1b domain shows different changes, when compared to the mechanism of the p67$^{phox}$-Rac1 complexes, but still characteristic for isolated Rac1b

domain. The p-loop and switch 1 show no significant change, for the same reasons as in p67$^{phox}$-Rac1 simulations, while switch 2 assumes a more open conformation but shows no unfolding. The flanking $\beta$-sheets show no significant change. Importantly, the 19-amino acid insertion shows significant contraction, similar to the one seen in the LOV2-Rac1b simulations as will be described below. The magnesium ion is complexed with the same partners as in the p67$^{phox}$-Rac1 inactive state. However, unlike the p67$^{phox}$-Rac1 active state, here the complexation partners change. The Thr17$_{Rac1}$ and Thr35$_{Rac1}$ residues are replaced with the Asp56$_{Rac1}$ residue, forming a complex geometry as in the isolated Rac1b active state.

The residues Ser22$_{Rac1}$, Thr25$_{Rac1}$, Asn26$_{Rac1}$, Gly30$_{Rac1}$ and Glu31$_{Rac1}$ are the Rac1b residues involved in the interaction between Rac1b and p67$^{phox}$. Interaction of Arg102$_{p67}$ with Ser22$_{Rac1}$ and Asn26$_{Rac1}$ shows no large change when comparing the two states. The connection between Asp67$_{p67}$ and Gly30$_{Rac1}$ stretches in the active state. Glu31$_{Rac1}$ residue changes orientation and gets more buried into the p67$^{phox}$ domain. The residues Ser37$_{p67}$, Asp67$_{p67}$, Arg102$_{p67}$, Asn104$_{p67}$ and Asp108$_{p67}$ are important residues involved in the p67$^{phox}$-Rac1b interaction surface.

The center of mass distances in the Rac1b complexes, both inactive and active states, behave similarly to the p67$^{phox}$-Rac1 inactive state. The changes induced in the p67$^{phox}$-Rac1 structure through Rac1 activation induce a separation of the two units, which is counter-intuitive.

The 19-amino acid insertion, in the inactive state, stabilizes the protein-protein interface by locking the switch 1 conformation through the adjacent $\alpha$1 contacts not directly involved in the interface. Specifically, the hydrogen bonds formed between Gly81$_{Rac1}$ and Thr24$_{Rac1}$ stabilize the protein-protein interface as well as the cintacts formed between Lys82$_{Rac1}$ and Tyr23$_{Rac1}$. However, upon activation, the 19-amino acid insertion contracts, breaking these connections, which induces structural changes in the $\alpha$5 helix adjacent to the $\beta$2 and $\beta$3 sheets. The part of the switch 1 region previously not involved in the protein-protein interface, with this, takes part in the interface stabilization. These structural changes cause the formation of new hydrogen bonds between p67$^{phox}$ and Rac1b, which are not present in the inactive state. Most prominent, a new connection is formed between Asn104$_{p67}$ and Gln181$_{Rac1}$ in addition to the hydrogen bond formed between Ala27$_{Rac1}$ and Gln181$_{Rac1}$. These new connections are locking two subunits further together and disallow the separation of the units, as seen in the p67$^{phox}$-Rac1 active state simulation. Due to the lacking of 19-amino acid insertion region in the p67$^{phox}$-Rac1 simulations, the activation of the Rac1 domain leads to a separation of the units as there is no additional stabilization of the units, as seen in the Rac1b simulation.

### 6.3.4 LOV2-Rac1b signaling mechanism

When comparing dark-inactive (dark state-LOV2-Rac1-GDP) to dark-active (dark state-LOV2-Rac1-GTP) states there are several typical changes induced in the

Rac1b domain that do not extend past the interface with LOV2. Switch 1 and switch 2 take an open, more flexible conformation while making a connection between them. The same has been seen in the isolated Rac1b domain. Furthermore, the p-loop shows no significant change. However, it does form a connection with the nucleotide in the active state. The $\beta$2 and $\beta$3 sheets show higher flexibility, breaking the connection with the $\alpha$5 helix, as described for the isolated Rac1b domain. Interestingly, the 19-amino acid insertion behaves as in the isolated Rac1b: upon activation it extends further, looping around the LOV2 domain. Our results suggest that the LOV2 domain sits on top of the 19-amino acid insertion and that this orientation and geometry of the 19-amino acid insertion in LOV2-Rac1b fusion protein acts as a 'loading arm'. This loading arm extends and contracts based on the activation state of the Rac1b domain, releasing or binding LOV2 to the Rac1b domain, respectively.

Structural changes of both LOV2 and Rac1b domains in all different activation states can also be seen in figures B.1 to B.4 presented in the appendix.

The magnesium ion changes its complexation partners going from dark-inactive to dark-active states. In the dark-inactive state the $Mg^{2+}$ ion is bound with $Thr17_{Rac1}$, the phosphate groups of GDP and two water molecules, while in the dark-active state those residues are replaced with $Thr35_{Rac1}$, $Asp56_{Rac1}$ and the GTP phosphate groups. It is interesting to note that there are no water molecules binding to the magnesium ion in the dark-active state.

Previous studies [206] suggest that the calcium ion present in the structure can have certain implications in the control of the activation mechanism. However, in our simulations we consistently see that the calcium ion is fully dissolved in the solvent and makes no connections to the protein. Hence, our simulations suggest that the calcium ion plays no role in stabilization of the protein construct in a calcium-deficient environment, as already described [206] for cytosol, the low-calcium environment.

The LOV2 domain shows no structural change while in the dark state. However, activation of the LOV2 domain induces significant structural changes in LOV2, some of which have previously been described [16]. Whereas in both dark state simulations (dark-inactive and dark-active) the J$\alpha$ helix remains stable and keeps the initial geometry, in the active state simulations we see that the geometry of the J$\alpha$ helix changes. It partially looses its helical structure but it also shows unwinding and breakage of its geometry. This is an important step in the signal transduction, as J$\alpha$ detachment and unfolding has been described as a key component in various LOV-systems. [16, 67–70] Furthermore, unfolding of the A'$\alpha$ helix has been shown [69] to be crucial step in the LOV-domain signal propagation. In our simulations, upon LOV2-activation, we see complete unfolding of the small A'$\alpha$ helix which does not occur in the dark states. Furthermore, activation of LOV2 induces significant torsions and coupling of the G$\beta$ and H$\beta$ sheets together with partial unfolding of the H$\beta$ sheet. This corresponds to early stages of the $\beta$-sheet tightening process, seen in

FTIR experiments on isolated LOV2 systems. [73, 214] Moreover, another key step in the LOV-signaling is the flip of the glutamine residue, adjacent to the FMN and conserved in most LOV domains. This Gln-flip has been described in the literature [16, 65, 66] as a key step in the signaling mechanism. In our simulations the Gln-flip occurs in all light state simulations, where the changes in the hydrogen-bonding network propagate through the LOV2 domain. The complete signaling pathway within LOV2-domain seems to be conserved, compared to the LOV-signaling described in the literature.

Activating the Rac1b domain, next to the already active LOV2 domain, induces several distinct changes in the structure of the fusion construct. Most notably, switches 1 and 2 take an open conformation, similar to those in the isolated Rac1b simulations, while the 19-amino acid insertion shows a slight extension towards LOV2. No significant changes are present in the p-loop and the flanking $\beta$ sheets.

The magnesium ion shows the same behavior as in the isolated Rac1b simulations. In the light-inactive state, magnesium is complexed with $Thr17_{Rac1}$, the phosphate groups from GDP and two water molecules. In the light-active state, however, $Thr17_{Rac1}$ is replaced with $Asp56_{Rac1}$ in addition to the GTP phosphate groups and a single water molecule. The $Ca^{2+}$ ion shows the same behavior as in the dark-inactive and dark-active simulations - it is fully dissolved in water, escaping the protein environment.

The separation of the LOV2 and Rac1b units is controlled by the activation state of the Rac1b domain. Based on our simulations, we propose that the 19-amino acid insertion together with the switch 2 region are responsible for the control of the unit binding.

Interestingly, activating the LOV2 domain, connected to an already active Rac1b domain causes several key structural changes. Most notably, unfolding of the J$\alpha$ helix causes significant contraction of the 19-amino acid insertion as well as a change of the conformation of switch 1 and switch 2, mediated by the $\beta$-sheet contacts. The conformation of switches 1 and 2 changes to an even more open conformation, further increasing the flexibility of the two regions. This is an important event, as it has previously been suggested [98] that the activation state of LOV2 may be the controlling step of the Rac1b's downstream effector binding while in the fusion construct. However, subsequent experimental studies [206] showed that the sensor-effector interface does not completely separate, as previously suggested. While the activation of the LOV2-domain may not be responsible for unit separation but rather the activation of Rac1b is controlling the association of the subunits. The LOV2 domain is controlling the flexibility and conformations of switches 1 and 2 and with them the binding of Rac1b to the downstream effector proteins. In the case of LOV2-Rac1b, the J$\alpha$ helix may have an opposite effect as seen in the LOV-TAP construct [215], in particular the unfolding of J$\alpha$ may not cause the release of the Rac1b unit from LOV2.

These results taken together suggest that the activation state of Rac1b is controlling several key structural aspects, from the unit binding to nucleotide hydrolysis[2], while the activation state of LOV2 further controls the partner binding as well as the open conformation of the switches 1 and 2 of Rac1b, responsible for the downstream effector binding.

## 6.4 Conclusions

The current study clearly demonstrates the early signaling mechanisms of Rac1 and Rac1b proteins and their differences. Furthermore, we suggest the importance and the effects each region of the Rac1 and Rac1b structures carries.

To further show that the signal transduction pathway seen in isolated Rac1 and Rac1b domains is a general Rac1 pathway, we analyzed the early signaling mechanisms of Rac1 and Rac1b protein constructs with RhoGDI, $p67^{phox}$ and LOV2. Not only do we consistently show the same signaling pathway within the Rac1 and Rac1b domains in these constructs, we were also able to further explain the structural changes induced by activation of one or both domains, their causes and effects that they have on the protein-protein interface and partner binding.

Our results show that upon activation of the Rac1 domain, switch 1 and switch 2 take a less-flexible, closed conformation, as opposed to Rac1b where the two regions are more flexible and are found in a released conformation, more open for downstream partner binding. We show that the 19-amino acid insertion in Rac1b is a key structural region in the control of signal propagation and opening of the binding sites. Moreover, we show the effects and structural changes induced in the vicinity of the binding sites and their potential effects on partner binding.

We also show the differences in partner binding and different geometrical restrictions required for binding with different protein partners, as in the cases of RhoGDI and $p67^{phox}$ domains with both Rac1 and Rac1b. Furthermore, we show the level of control each of the domains have in LOV2-Rac1b fusion construct.

However, important questions regarding the role that Rac1 and Rac1b play in control of downstream pathways, still need to be understood as well as their role in tumor development and progression. Also, a question of GDP/GTP exchange and Rac1 and Rac1b domain activation still needs to be fully elucidated.

[2]exchange of nucleotides, GTP by GDP

# 7 Material and Methods

## 7.1 Generation of RsLOV initial structures

Initial structures for all RsLOV simulations were generated based on the crystal structure of *R. sphaeroides* LOV domain in the dark-adapted state (PDB code: 4HJ6) [52]. All crystallographic waters were discarded and the structures were solvated using GROMACS built-in tools. To create the starting structures for our light-state wild type RsLOV simulations, we generated a FMN-cysteinyl adduct by forming a covalent bond between the FMN-C4a and Cys-S atoms. The initial structures for the mutant simulations were generated by introducing the corresponding single-point mutations to the structures of dark- and light-state wildtype RsLOV using the Modeller v9.11 software package, [101] employing the homology modeling method. [216, 217] The symmetrical dark-dark dimer initial structure has been generated by in-place exchange of the unit B with a copy of unit A.

## 7.2 Generation of CrLOV initial structures

For protein structure predictions the MODELLER version 9.11 software package [101] was used, employing the homology modelling method (comparative modelling). The model of each mutant has been generated based on the crystal structure of native CrLOV1 (PDB code: 1N9L, determined by Fedorov et al. [58]), by introducing the required point mutations: F41Y, F41Y C57A, F41Y C57S and F41Y C57G.

## 7.3 Generation of Rac1 and Rac1b initial structures

The crystal structure of human wild type Rac1 protein (PDB code: 3TH5) [180] has been used as a model template for generating the starting structures for our Rac1-GDP and Rac1-GTP simulations. The crystal structure of Rac1b (PDB code: 1RYF) [175] has been used to model the initial structures for Rac1b-GDP and Rac1b-GTP simulations. Secondary structures of Rac1 and Rac1b are visualized in figure 6.1.2 with secondary structural elements labeled for both structures.

Initial structures for our RhoGDI-Rac1-GDP and RhoGDI-Rac1-GTP simulations have been generated based on the crystal structure of the Rac1-RhoGDI complex

(PDB code: 1HH4). [207] Active and inactive initial structures of p67$^{phox}$-Rac1 as well as p67$^{phox}$-Rac1b have been modeled based on the crystal structure of the Rac1-p67$^{phox}$ complex (PDB code: 1e96). [201] To simulate the behavior of photoactivatable Rac1b-GDP and Rac1b-GTP, containing a wild type LOV2 domain in either the dark or the light state, the initial structures were generated based on the wild type LOV2-Rac1 crystal structure (PDB code: 2WKP) [98].

In each of the modeled structures, missing residues as well as GDP or GTP residues have been modeled using the Modeller v9.11 software package [101], employing the homology modeling method (comparative modeling). [216, 217]

## 7.4 Simulation details

To resolve the behavior of the different protein systems at a molecular level, we generated MD trajectories with the GROMACS version 4.6.5 software package using the GROMOS43a1 forcefield. [102] We have used this forcefield for both the free-state and the adduct-state flavin. It has been shown on numerous protein-solvent systems that this forcefield reliably reproduces experimental results. [103, 104] In a recent study [104], the GROMOS43a1 forcefield has been compared to the CHARMM27, AMBER03, OPLS and GROMOS53a6 forcefields. This study analyzed the effects of each of the forcefields on the conformational development of insulin over the simulation time and compared the results with the experimentally observed insulin behavior. While all forcefields favor different protein conformations and evolution over time, the study shows that GROMOS43a1 together with CHARMM27 best describe the available experimental data. Full particle-mesh-Ewald (PME) electrostatics was used for all of our simulations with a real-space cutoff of 0.9 nm, and van der Waals interactions were computed with a cutoff of 1.4 nm. Proteins were placed in a cubic box where the distance between the edge of the box and the closest protein atom was set to at least 1.0 nm. To neutralize each of the systems, an adequate number of sodium or chloride ions were added to each simulation box at random positions, depending on the system investigated. For each system, an equilibration run was carried out for 5 ns using the NPT ensemble at 300 K and 1 bar, with all parts of the system coupled to the Nosé-Hoover-thermostat and the Parrinello-Rahman barostat. Following the equilibration phase, production runs of 20 ns in CrLOV simulations, 50 ns for RsLOV system and 100 ns of Rac1 and Rac1b systems and their constructs (50 ns in the cases of RhoGDI and p67$^{phox}$ constructs) of noninvasive thermostatting [218] was performed, where both the protein and the cofactors were decoupled from the thermostat, while the solvent and ions remained coupled. This non-invasive method allows the protein to follow its natural dynamics by being thermostatted through solvent contacts. It also allows the protein to sample configurations which are far from equilibrium. For numerical integration of the equations of motion we have used the Leap-frog integrator with a timestep of 2 fs in CrLOV and RsLOV simulations, while in the Rac1 and Rac1b simulations

we used a timestep of 1 fs. All bond lengths were constrained using LINCS algorithm. [219] Parameters for FMN are available in the GROMOS force field, while the parameters of Neiss and Saalfrank [220] were used to describe the behavior of the FMN-photoadduct.

## 7.5 Electrostatics calculations

To study the protein-protein interactions we used the Adaptive Poisson-Boltzmann Solver (APBS) [221] plugin for PyMOL [222] to calculate the electrostatic potentials by solving the Linearized Poisson-Boltzmann equation (LPBE). The Poisson-Boltzmann equation and its linearized form allow calculations of the electrostatic effects and interactions of biomolecular systems in a water environment. The input structures for APBS calculations were taken from MD simulations. The file-format of the protein structures was converted from PDB to PQR using the PDB2PQR web server [223, 224] with the CHARMM27 [225] forcefield. All structures for the present ligands, containing all the necessary bonding information, were obtained using the PRODRG [226] web server in MOL2 format.

## 7.6 Calculation of pKa values of titratable amino acids

The protonation state for each titratable amino-acid was estimated using the finite difference Poisson-Boltzmann (FDPB) continuum electrostatics methodology as implemented in H++ web server v3.2, available at http://biophysics.cs.vt.edu/. [114] The equilibrated simulation structures, generated as described above, were used as input structures for pKa calculations. Calculations were performed by using the following physical conditions: salinity was set to 0.15 and the internal and external dielectric constants were set to 10 and 80, respectively. The input structures were protonated assuming pH of 7.9. Correct orientation of His, Gln and Asn groups and the correct assignment of the protonation position of His residues were based on van der Waals contacts and hydrogen bonding.

## 7.7 Secondary structure assignment

Assignment of secondary structures have been performed using the DSSP (Define Secondary Structure of Proteins) program, [227, 228] employing the inbuilt GROMACS interface. The DSSP algorithm calculates and assigns the most likely secondary structure for each residue in the 3D structure of the protein. [227] This is done based on a database of secondary structures computationally derived from

the PDB database. DSSP calculates the electrostatic energy of all identified intra-backbone hydrogen bonds. Based on this parameter the secondary structure is assigned as one of the eight types, $3_{10}$ helix, $\alpha$ helix, $\pi$ helix, $\beta$ bridge, $\beta$ bulge, turn, loop, and blank (sequence break).

# 8 Summary

Light-, oxygen- and voltage- (LOV) sensing domains are blue-light sensitive proteins with a number of biological functions in various organisms. The active state is generated through reversible formation of a covalent bond between FMN and an adjacent cystein. This active state thermally recovers back to the dark state, with time. Light-triggered signal transduction from the LOV-core to the effector domain or binding partner occurs through the flanking terminal helices. RsLOV is a small LOV domain that has a helix-turn-helix motif as an effector region. In this work we present the early signaling mechanism, subunit behavior and association properties of the RsLOV and of the CrLOV domains upon the light-induced activation.

We start by proposing the mechanism of early-signal propagation in RsLOV upon adduct formation. After the perturbation of the FMN-binding pocket, signal is transfered over the anti-parallel $\beta$-sheets to the A'$\alpha$ helix, inducing loss of the helical content of the A'$\alpha$ helix. As the A'$\alpha$ helix is in direct contact with the J$\alpha$ helix, signal is directly transduced to the helix-turn-helix motif, formed by the J$\alpha$ and K$\alpha$ flanking helices. We also show that the signal is transduced to the second unit of the RsLOV dimer through the dimerization surface, and the changes induced in the helical bundle by the signal propagation. We further propose an alternative, dark-state conformation present in the second unit. Finally, we suggest a mutual stabilization exhibited by the two subunits as well as the association properties of the two units, forming the RsLOV.

To study the signal transduction dynamics in CrLOV, we introduced a number of single-point mutations in CrLOV with a Tyr nearby the flavine. This introduced Tyr abolishes the adduct formation and changes the mechanism from adduct formation to an electron transfer (ET). Our results suggest a proton coupled electron transfer (PCET) occuring in F41Y mutant, even when the conserved Cys is preserved. The introduced Tyr acts as an electron donor while FMN acts as a terminal acceptor in the CrLOV ET chain. The MD simulations reveal that F41Y mutation has only minor influence on the overall structure, with small perturbations occurring around Tyr41. The hydrogen bonding network built around FMN, including the two Asn residues, remains intact. The Tyr41 residue can form a hydrogen bond with FMN, however the direct interaction with FMN-N5 is highly unlikely. Additionally, MD simulations reveal different accessibility of solvent molecules into the binding pocket, thus supporting different pH dependence of the individual mutations.

ET processes are crucial components of a wide range of protein signaling mechanisms. Understanding the complete ET pathway is thus of key importance in the

study of ET proteins. We have developed novel, fully-parameterized algorithms to predict important residues involved in electron transfer pathways in proteins. Based on the structural information of a protein, we are able to calculate the rates of possible ET transfer jumps and predict the dominant ET pathway chain. We are also able to predict the residues that are most likely the initial donor or the terminal acceptor of the ET chain. Here, we describe the complete set of algorithms and the methods as implemented in the etFIND software package. The program has been successfully tested on a set of 10 proteins for which the ET pathways have been previously fully described. Finally, we apply etFIND in understanding and guiding the site directed mutagenesis to alter ET chains in proteins, on an example of DNA Photolyase from *E. Coli*. We applied etFIND on CrLOV WT and its single-point mutants, described above. In all cases etFIND predicts Tyr41 to be the donor and FMN to be the acceptor of ET. It also predicts a direct ET between the two residues, with no intermediate charge carriers involved.

Rac1 is a fast cycling small GTPase that cycles between an inactive GDP-bound and an active GTP-bound states. Rac1 has been linked with regulation of cell motility and growth control as well as with other cellular processes, phagocytosis, adhesion, cell proliferation and differentiation. It has been linked with malignancy as well as with multiple renal and cardiac diseases. A spliced variant of Rac1, called Rac1b, has 19-amino acids inserted which drastically change the behavior of this protein. Rac1b has a very high guanine nucleotide cycling rate and decreased GTPase activity. It is up-regulated in various human tumors and it has been linked with malignant progressions of a number of carcinomas. It is thus of great importance to fully understand how structural differences between the two isoforms guide changes in their function. Here we report on MD simulations that yield insight into the early signaling mechanisms of the two protein variants and their differences.

Rac1, compared to Rac1b, shows less flexible, closed conformations of switches 1 and 2, around the binding site. Additionally, the flanking $\beta$-sheets show structural rearrangements needed to accommodate the changes following the domain activation upon GDP/GTP exchange. Contrary to Rac1, in Rac1b the entire nucleotide binding region has a more flexible, open conformation of the switch 1 and switch 2 regions. Moreover, we show that the 19-amino acid insertion plays a key role in structural control of the Rac1b domain and its binding sites, upon nucleotide exchange in Rac1b. We further studied the signaling mechanism of the protein constructs of Rac1 and Rac1b with their protein partners, RhoGDI and $p67^{phox}$. The mechanism of Rac1 and Rac1b is conserved in their $p67^{phox}$ and RhoGDI complexes.

Finally, we show the early signal transduction dynamics of the LOV2-Rac1b fusion construct. We further show the important aspects of the LOV2-Rac1b structure and the structural effects that the 19-amino acid insertion has. The LOV2 domain sits on top of the 19-amino acid insert which acts as a loading arm. This loading arm contracts or extends, bringing the two subunits closer together or separating them from each other. While we observe the importance of the magnesium ion in nucleotide binding and recognition, the calcium ion, in our simulations, shows no

regulative role in a calcium-deficient environment. The LOV2 domain, like Rac1b, also preserves its molecular mechanism. In particular, the effects on the J$\alpha$ and A'$\alpha$ helices along with the coupling of the $\beta$-sheets are conserved. We propose that Rac1b has control on unit binding and nucleotide hydrolysis. On the other hand, the LOV domain controls partner binding through the release of the binding regions of Rac1b, responsible for downstream effector recognition and binding.

# Appendix

## A Analyses of secondary structure changes of RsLOV trajectories

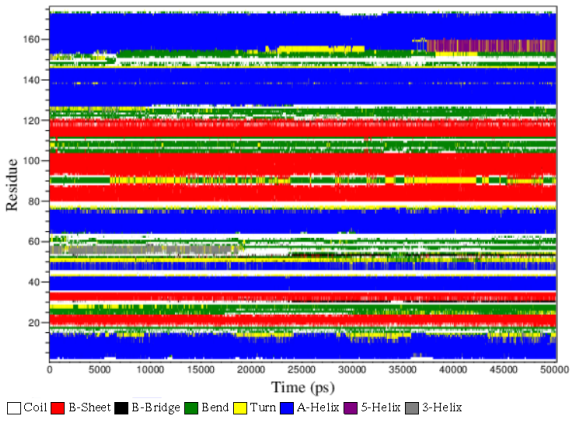


**Figure A.1** Analysis of the secondary structure changes during the dark state monomer trajectory.

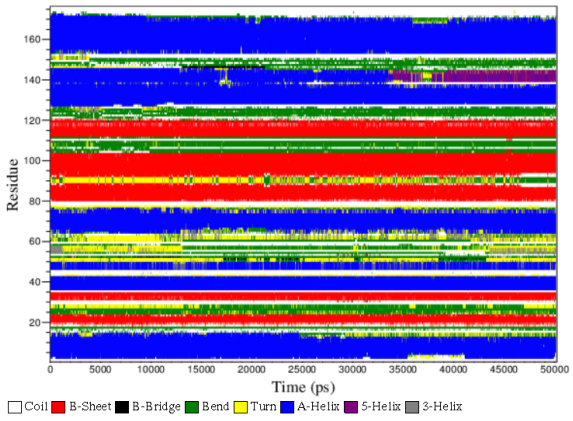


**Figure A.2** Analysis of the secondary structure changes during the light state monomer trajectory.

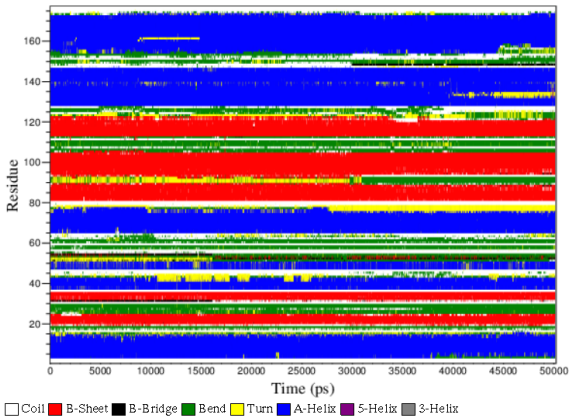


**Figure A.3** Analysis of the secondary structure changes during the trajectory of the dark state monomer, taken after light-dark simulation has completed.

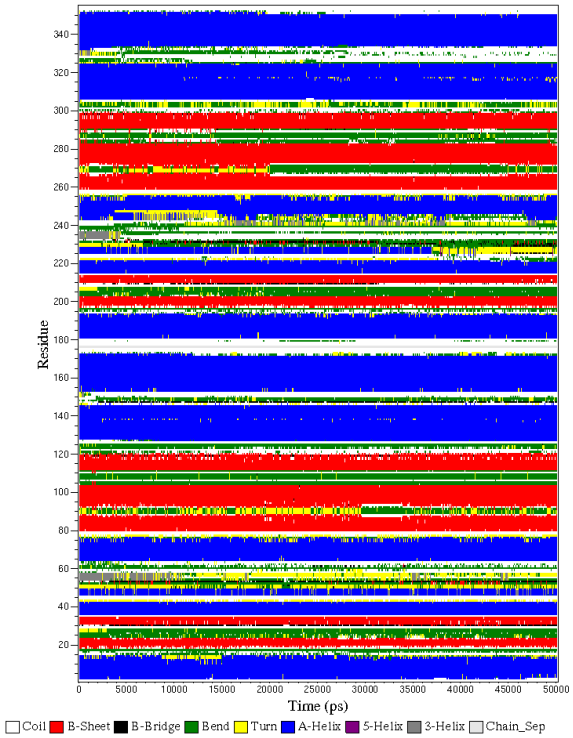


**Figure A.4** Analysis of the secondary structure changes during the dark-dark dimer trajectory.

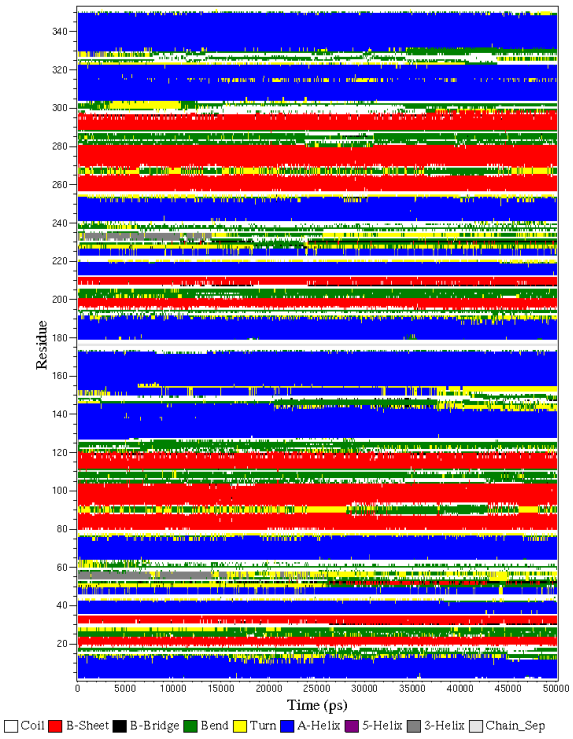


**Figure A.5** Analysis of the secondary structure changes during the light-dark dimer trajectory.

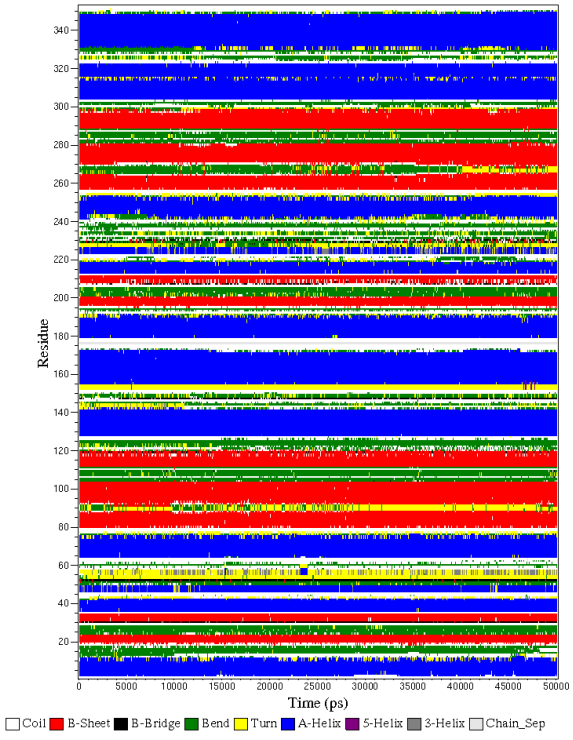


**Figure A.6** Analysis of the secondary structure changes during the light-light dimer trajectory.

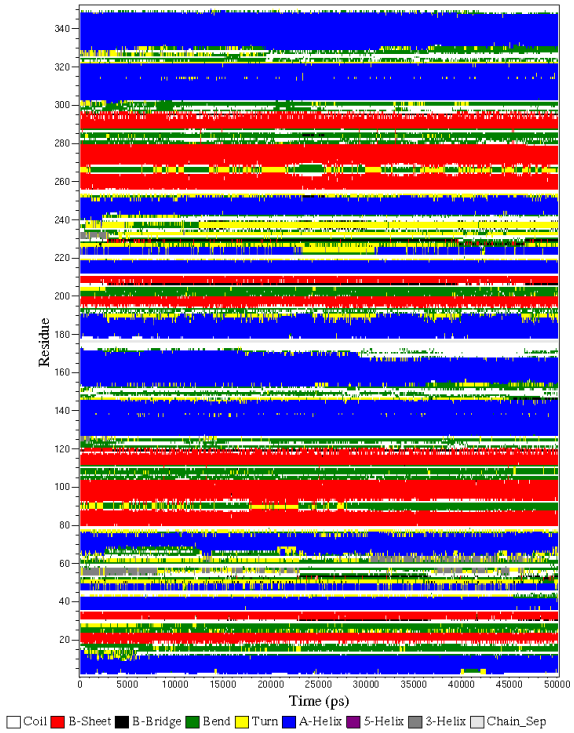


**Figure A.7** Analysis of the secondary structure changes during the symmetrical dark-dark dimer trajectory.

# B Analyses of secondary structure changes of LOV2-Rac1b trajectories

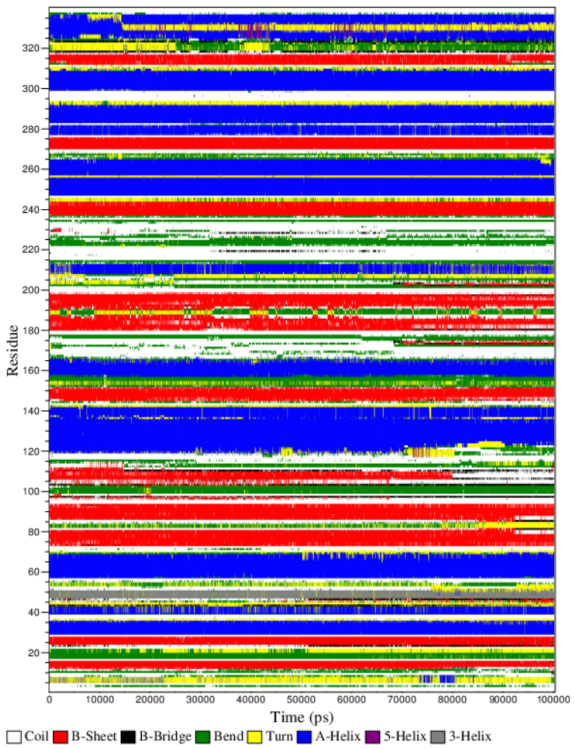


**Figure B.1** Analysis of the secondary structure changes during the trajectory where both LOV2 and Rac1b domains are in their respective inactivated states.

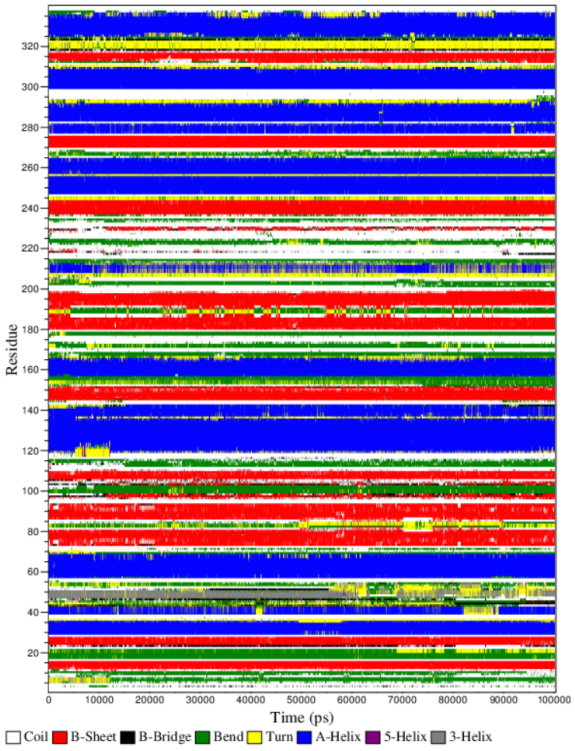


**Figure B.2** Analysis of the secondary structure changes during the trajectory where LOV2 domain is in a dark state while Rac1b domain is in an active state.

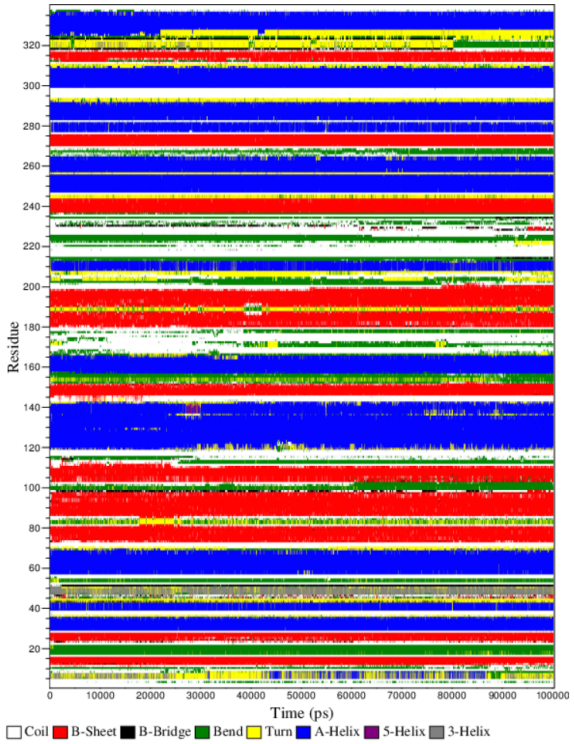


**Figure B.3** Analysis of the secondary structure changes during the trajectory where LOV2 domain is in an active state while Rac1b domain is in a inactive state.

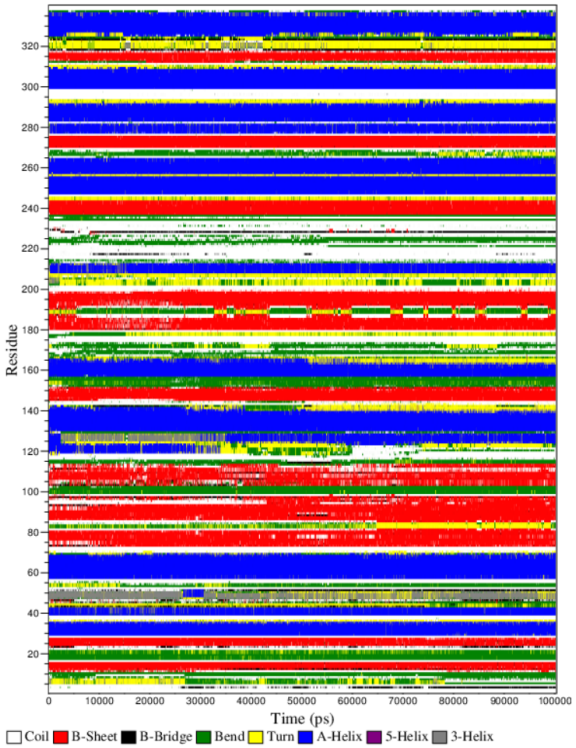


**Figure B.4** Analysis of the secondary structure changes during the trajectory where both LOV2 and Rac1b domains are in their respective activated states.

# C Supporting Information to: Switching from adduct formation to electron transfer in a light-oxygen-voltage domain containing the reactive cysteine

## C.1 Mutagenesis

The gene encoding for the cDNA of *Cr*LOV1 wt was used as a template for site-directed mutagenesis. PCR was performed using *Pfu* DNA polymerase (Thermo Scientific, Waltham, MA, US) and the primers given in table C.1. The sequence of the plasmids carrying the desired nucleotide substitutions was confirmed by DNA sequencing (Seqlab Sequence Laboratories, Göttingen, Germany)

## C.2 UV/Vis spectra of the different *Cr*LOV1-F41Y mutants

UV/Vis spectra of the *Cr*LOV1-F41Y mutants were recorded in the range of 300-600 nm using a Lambda 9a spectrometer (Perkin Elmer, Waltham, MA, US). Spectra were recorded before and after transient absorption measurements in order to control the stability of the samples.
Figure C.1 shows the mutants *Cr*LOV1-F41Y, *Cr*LOV1-F41Y/C57S, *Cr*LOV1-F41Y/C57A and *Cr*LOV1-F41Y/C57G in comparison with the wild type. All mutants exhibit the typical fine structured absorption bands around 445 nm and 360 nm, showing only minor wavelength shifts compared to the wild type.

**Table C.1** Oligonucleotide primers used for site-directed mutagenesis

| Primer name | DNA base sequence (5' to 3') |
|---|---|
| CrLOV1-F41Y-s [a] | CAGCGAGGGGTATTa[b]TGCCATGAC |
| CrLOV1-F41Y-as [a] | GTCATGGCAtAATACCCCTCGCTG |
| CrLOV1-C57S-s | GTGCTTGGTCACAACaGCCGCTTCCT |
| CrLOV1-C57S-as | AGGAAGCGGCtGTTGTGACCAAGCAC |
| CrLOV1-C57A-s | CTTGTTCACAACgcgCGCTTCCTC |
| CrLOV1-C57A-as | GAGGAAGCGcgcGTTGTGACCAAG |
| CrLOV1-C57G-s | gGCCGCTTCCTCCAAGGCGAGGGCACGGACCCCAA |
| CrLOV1-C57G-as | CTTGGAGGAAGCGGCcGTTGTGACCAAGCACCTCATCGG |

[a] s: sense primer; as: antisense primer [b] The mutated DNA bases in the specific codons are highlighted by small letters.

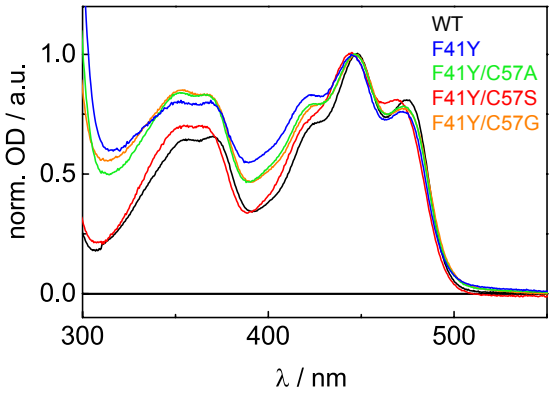


**Figure C.1** UV/Vis spectra of the *Cr*LOV1-F41Y mutants measured in the range of 300-600 nm.

**Table C.2** Absorption maxima of the $S_0 \rightarrow S_1$ and $S_0 \rightarrow S_2$ transitions of the *Cr*LOV1 mutants.

| | $S_0 \rightarrow S_1$ (nm) | $S_0 \rightarrow S_2$ (nm) |
|---|---|---|
| F41Y | 422/445/473 | 368/352 |
| F41Y/C57A | 424/446/473 | 367/352 |
| F41Y/C57S | 422/444/470 | 365/354 |
| F41Y/C57G | 425/447/474 | 368/354 |

## C.3 Interpretation of the $D_2$ of *Cr*LOV1-F41Y/C57S and *Cr*LOV1-F41Y/C57A

A contribution of $FMNH^{\bullet}$ and the corresponding radical of the electron donor Tyr were assumed to be neccessary to describe the DADS associated with the slower decaying species of *Cr*LOV1-F41Y/C57S and F41Y/C57A. In the case of F41Y/C57S the linear combination of $FMNH^{\bullet}$, FMNox and $TyrO^{\bullet}$ was not sufficient to describe $D_2$ in figure C.2A adequately. Taking into account a contribution of the FMN radical anion, $FMN^{\bullet-}$, results in the red line in figure C.2A and is in good agreement with the DADS.
In contrast, $D_2$ of *Cr*LOV1-F41Y/C57A (figure C.2B) could be best described using a linear combination of $FMNH^{\bullet}$ plus $FMN^{\bullet-}$ minus FMNox. In this case, the contribution of the FMN radical anion is larger and the contribution of $TyrO^{\bullet}$ did not improve the spectral fitting. Thus, we can not prove unequivocably that $TyrO^{\bullet}$ is stabilized as the counter-radical of $FMNH^{\bullet}$ in *Cr*LOV1-F41Y/C57A.

## C.4 DADS of *Cr*LOV1-F41Y/C57G

The transient absorption measurement of *Cr*LOV1-F41Y/C57G was performed on a 100 $\mu$s streak window with laser excitation at 447 nm. Global lifetime analysis of the resulting 2D TA data matrix using two exponential functions results in the DADS

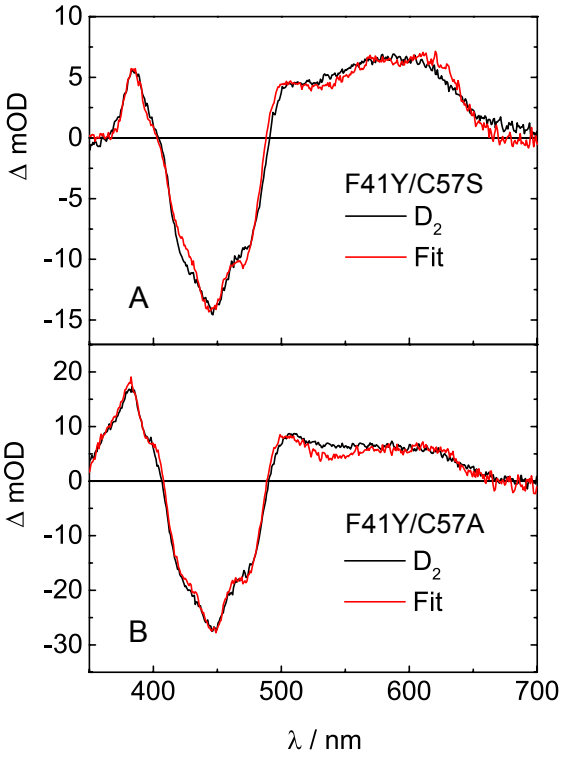


**Figure C.2 A:** $D_2$ of *Cr*LOV1-F41Y/C57S (black line) obtained by global lifetime analysis. The spectrum could be described in good agreement by using a linear combination (fit, red line) of $FMNH^{\bullet}$, $FMN^{\bullet-}$, FMNox and $TyrO^{\bullet}$. **B:** $D_2$ of *Cr*LOV1-F41Y/C57A (black line) could be best described by using a linear combination (fit, red line) of $FMNH^{\bullet}$, $FMN^{\bullet-}$ and FMNox. In this case, inclusion of $TyrO^{\bullet}$ did not improve the result.

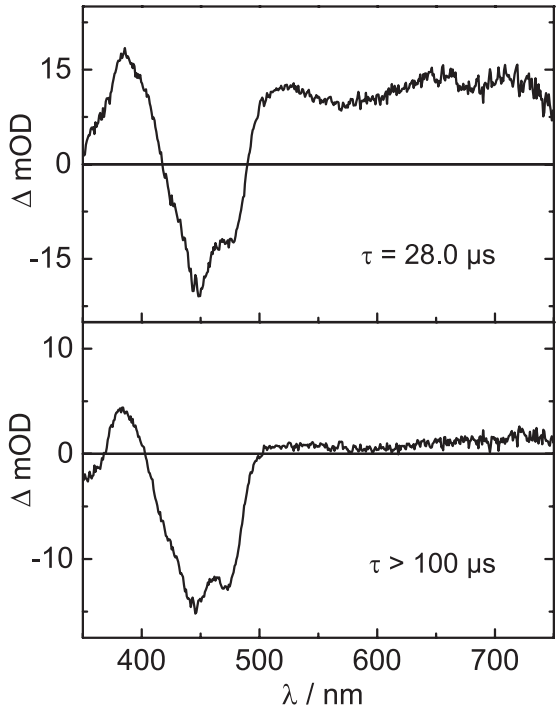


**Figure C.3** DADS of *Cr*LOV1-F41Y/C57G obtained by global lifetime analysis. **A:** $D_1$ decaying with a rate constant of 28.0 $\mu$s **B:** $D_2$ non-decaying within 100.0 $\mu$s.

shown in figure C.3. Both DADS exhibit characteristic features of the FMN triplet state indicating a biphasic triplet decay like already observed for *Cr*LOV1-C57G [12]. Further details are described in the main text of this publication. However it is noteworthy, that the ground state bleach of FMNox in $D_2$ is not recovering proportionally to the decrease of positive absorption signals in comparison with $D_1$. This is due to irreversible photo-damage of some fraction of *Cr*LOV1-F41Y/C57G during the TA measurement.

## C.5 Dihedral angle distribution between F41Y and FMN

Dihedral angle distributions between Tyr41-CE2, FMN-(C5a-C4a) and Tyr41-CE1 of all *Cr*LOV1-F41Y mutants obtained from analyzing the 20 ns MD trajectories. The distribution curves are broad in all cases with varying maxima from 18 ° (F41Y/C57S) to 22.5 ° (F41Y). CrLOV1-F41Y/C57A additionally has a second minor peak. The dihedral angle distribution indicate a considerable flexibility between Tyr41 and FMN possibly leading to different geometry-dependent reaction pathways.

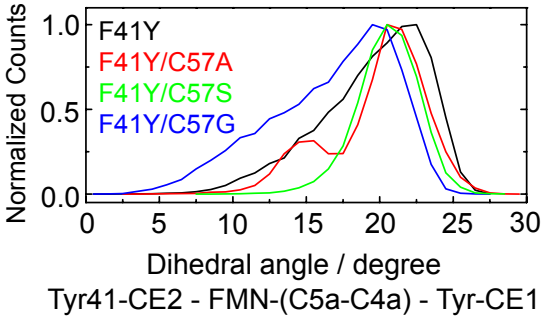


**Figure C.4** Dihedral angle distributions between Tyr41-CE2, FMN-(C5a-C4a) and Tyr41-CE1 of all *Cr*LOV1-F41Y mutants.

# Bibliography


[1] J. Downward: *The ins and outs of signalling*. Nature **411**, 759 (2001).

[2] B. F. Volkman, D. Lipson, D. E. Wemmer *et al.*: *Two-State Allosteric Behavior in a Single-Domain Signaling Protein*. Science **291**, 2429 LP (2001).

[3] E. B. Purcell, D. Siegal-Gaskins, D. C. Rawling *et al.*: *A photosensory two-component system regulates bacterial cell attachment*. Proc. Natl. Acad. Sci. USA **104** (2007).

[4] T. E. Swartz, S. B. Corchnoy, J. M. Christie *et al.*: *The Photocycle of a Flavin-binding Domain of the Blue Light Photoreceptor Phototropin*. J. Biol. Chem. **276**, 36493 (2001).

[5] J. Purschwitz, S. Müller, C. Kastner *et al.*: *Seeing the rainbow: light sensing in fungi*. Curr. Opin. Microbiol. **9**, 566 (2006).

[6] E. Huala, P. W. Oeller, E. Liscum *et al.*: *Arabidopsis NPH1: A Protein Kinase with a Putative Redox-Sensing Domain*. Science **278**, 2120 LP (1997).

[7] S. Crosson und K. Moffat: *Structure of a flavin-binding plant photoreceptor domain: Insights into light-mediated signal transduction*. Proc. Natl. Acad. Sci. USA **98**, 2995 (2001).

[8] S. Crosson, S. Rajagopal, und K. Moffat: *The LOV Domain Family: Photoresponsive Signaling Modules Coupled to Diverse Output Domains*. Biochemistry **42**, 2 (2003).

[9] A. Losi: *Flavin-based Blue-light Photosensors: A Photobiophysics Update*. Photochem. Photobiol. **83**, 1283 (2007).

[10] J. T. M. Kennis, S. Crosson, M. Gauden *et al.*: *Primary Reactions of the LOV2 Domain of Phototropin, a Plant Blue-Light Photoreceptor*. Biochemistry **42**, 3385 (2003).

[11] C. Bauer, C.-R. Rabl, J. Heberle *et al.*: *Indication for a Radical Intermediate Preceding the Signaling State in the LOV Domain Photocycle*. Photochem. Photobiol. **87**, 548 (2011).

[12] R. J. Kutta, K. Magerl, U. Kensy *et al.*: *A search for radical intermediates in the photocycle of LOV domains*. Photochem. Photobiol. Sci. **14**, 288 (2015).

[13] A. Losi, T. Kottke, und P. Hegemann: *Recording of Blue Light-Induced Energy and Volume Changes within the Wild-Type and Mutated Phot-LOV1 Domain from Chlamydomonas reinhardtii*. Biophys. J. **86**, 1051 (2004).

[14] W. Gong, B. Hao, S. S. Mansy *et al.*: *Structure of a biological oxygen sensor: A new mechanism for heme-driven signal transduction.* Proc. Natl. Acad. Sci. USA **95**, 15177 (1998).

[15] Y.-Z. Gu, J. B. Hogenesch, und C. A. Bradfield: *The PAS Superfamily: Sensors of Environmental and Developmental Signals.* Annu. Rev. Pharmacol. Toxicol. **40**, 519 (2000).

[16] E. Peter, B. Dick, und S. A. Baeurle: *Mechanism of signal transduction of the LOV2-Jalpha photosensor from Avena sativa.* Nat. Commun. **1**, 122 (2010).

[17] A. T. Vaidya, C.-H. Chen, J. C. Dunlap *et al.*: *Structure of a Light-Activated LOV Protein Dimer That Regulates Transcription.* Sci. Signal. **4**, ra50 LP (2011).

[18] J. Blumberger: *Recent Advances in the Theory and Molecular Simulation of Biological Electron Transfer Reactions.* Chem. Rev. **115**, 11191 (2015).

[19] C. C. Moser und P. L. Dutton: *Engineering protein structure for electron transfer function in photosynthetic reaction centers.* Biochim. Biophys. Acta - Bioenerg. **1101**, 171 (1992).

[20] R. A. Marcus: *On the Theory of Oxidation-Reduction Reactions Involving Electron Transfer. I.* J. Chem. Phys. **24**, 966 (1956).

[21] R. A. Marcus: *Electrostatic Free Energy and Other Properties of States Having Nonequilibrium Polarization. I.* J. Chem. Phys. **24**, 979 (1956).

[22] J. J. Hopfield: *Electron Transfer Between Biological Molecules by Thermally Activated Tunneling.* Proc. Natl. Acad. Sci. USA **71**, 3640 (1974).

[23] C. C. Page, C. C. Moser, X. Chen *et al.*: *Natural engineering principles of electron tunnelling in biological oxidation-reduction.* Nature **402**, 47 (1999).

[24] D. N. Beratan, J. N. Onuchic, und J. J. Hopfield: *Electron tunneling through covalent and noncovalent pathways in proteins.* J. Chem. Phys. **86**, 4488 (1987).

[25] D. N. Beratan, J. N. Betts, und J. N. Onuchic: *Protein electron transfer rates set by the bridging secondary and tertiary structure.* Science **252**, 1285 (1991).

[26] T. R. Prytkova, I. V. Kurnikov, und D. N. Beratan: *Coupling Coherence Distinguishes Structure Sensitivity in Protein Electron Transfer*. Science **315**, 622 LP (2007).

[27] D. N. Beratan und I. A. Balabin: *Heme–copper oxidases use tunneling pathways.* Proc. Natl. Acad. Sci. USA **105**, 403 (2008).

[28] C. C. Moser, C. C. Page, und P. L. Dutton: *Darwin at the molecular scale: selection and variance in electron tunnelling proteins including cytochrome c oxidase.* Philos. Trans. R. Soc. B Biol. Sci. **361**, 1295 LP (2006).

[29] C. C. Moser, S. E. Chobot, C. C. Page *et al.*: *Distance metrics for heme protein electron tunneling*. Biochim. Biophys. Acta - Bioenerg. **1777**, 1032 (2008).

[30] B. Stewart und C. Wild: *World cancer report 2014*. International Agency for Research on Cancer, Geneva, Switzerland, WHO Press (2014).

[31] T. R. Golub, D. K. Slonim, P. Tamayo *et al.*: *Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring*. Science **286**, 531 LP (1999).

[32] J. E. Visvader: *Cells of origin in cancer*. Nature **469**, 314 (2011).

[33] U. Alon, N. Barkai, D. A. Notterman *et al.*: *Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays*. Proc. Natl. Acad. Sci. USA **96**, 6745 (1999).

[34] S. Peng, Q. Xu, X. B. Ling *et al.*: *Molecular classification of cancer types from microarray data using the combination of genetic algorithms and support vector machines*. FEBS Lett. **555**, 358 (2003).

[35] D. Hanahan und R. A. Weinberg: *Hallmarks of Cancer: The Next Generation*. Cell **144**, 646 (2011).

[36] D. Hanahan und R. A. Weinberg: *The Hallmarks of Cancer*. Cell **100**, 57 (2000).

[37] F. Michor, Y. Iwasa, und M. A. Nowak: *Dynamics of cancer progression*. Nat. Rev. Cancer **4**, 197 (2004).

[38] R. S. Kerbel: *Expression of multi-cytokine resistance and multi-growth factor independence in advanced stage metastatic cancer. Malignant melanoma as a paradigm*. Am. J. Pathol. **141**, 519 (1992).

[39] K. Rajalingam, R. Schreck, U. R. Rapp *et al.*: *Ras oncogenes and their downstream targets*. Biochim. Biophys. Acta - Mol. Cell Res. **1773**, 1177 (2007).

[40] M. Takata, Y. Goto, N. Ichii *et al.*: *Constitutive Activation of the Mitogen-Activated Protein Kinase Signaling Pathway in Acral Melanomas*. J. Invest. Dermatol. **125**, 318 (2005).

[41] D. A. Altomare und J. R. Testa: *Perturbations of the AKT signaling pathway in human cancer*. Oncogene **24**, 7455 (2005).

[42] M. S. Song, L. Salmena, und P. P. Pandolfi: *The functions and regulation of the PTEN tumour suppressor*. Nat. Rev. Mol. Cell Biol. **13**, 283 (2012).

[43] S. W. Lowe und A. W. Lin: *Apoptosis in cancer*. Carcinogenesis **21**, 485 (2000).

[44] D. Hanahan und L. M. Coussens: *Accessories to the Crime: Functions of Cells Recruited to the Tumor Microenvironment*. Cancer Cell **21**, 309 (2012).

[45] A. Eger und W. Mikulits: *Models of epithelial–mesenchymal transition*. Drug Discov. Today Dis. Model. **2**, 57 (2005).

[46] T. T. Onder, P. B. Gupta, S. A. Mani *et al.*: *Loss of E-Cadherin Promotes Metastasis via Multiple Downstream Transcriptional Pathways.* Cancer Res. **68**, 3645 LP (2008).

[47] P. Hegemann: *Algal Sensory Photoreceptors.* Annu. Rev. Plant Biol. **59**, 167 (2008).

[48] S. Anderson, V. Dragnea, S. Masuda *et al.*: *Structure of a Novel Photoreceptor, the BLUF Domain of AppA from Rhodobacter sphaeroides.* Biochemistry **44**, 7998 (2005).

[49] S. Endres, J. Granzin, F. Circolone *et al.*: *Structure and function of a short LOV protein from the marine phototrophic bacterium Dinoroseobacter shibae.* BMC Microbiol. **15**, 30 (2015).

[50] M. Salomon, J. M. Christie, E. Knieb *et al.*: *Photochemical and mutational analysis of the FMN-binding domains of the plant blue light receptor, phototropin.* Biochemistry **39** (2000).

[51] P. L. Freddolino, K. H. Gardner, und K. Schulten: *Signaling mechanisms of LOV domains: new insights from molecular dynamics studies.* Photochem. Photobiol. Sci. **12**, 1158 (2013).

[52] K. S. Conrad, A. M. Bilwes, und B. R. Crane: *Light-Induced Subunit Dissociation by a Light–Oxygen–Voltage Domain Photoreceptor from Rhodobacter sphaeroides.* Biochemistry **52**, 378 (2013).

[53] S. Metz, A. Jager, und G. Klug: *Role of a short light, oxygen, voltage (LOV) domain protein in blue light- and singlet oxygen-dependent gene regulation in Rhodobacter sphaeroides.* Microbiology **158** (2012).

[54] A. I. Nash, R. McNulty, M. E. Shillito *et al.*: *Structural basis of photosensitivity in a bacterial light-oxygen-voltage/helix-turn-helix (LOV-HTH) DNA-binding protein.* Proc. Natl. Acad. Sci. USA **108** (2011).

[55] B. D. Zoltowski und B. R. Crane: *Light activation of the LOV protein vivid generates a rapidly exchanging dimer.* Biochemistry US **47** (2008).

[56] A. Losi. *LOV Proteins: Photobiophysics*, 1312–1316. Springer Berlin Heidelberg, Berlin, Heidelberg (2013).

[57] S. Crosson und K. Moffat: *Photoexcited structure of a plant photoreceptor domain reveals a light-driven molecular switch.* Plant Cell **14** (2002).

[58] R. Fedorov, I. Schlichting, E. Hartmann *et al.*: *Crystal structures and molecular mechanism of a light-induced signaling switch: The Phot-LOV1 domain from Chlamydomonas reinhardtii.* Biophys. J. **84** (2003).

[59] D. Nozaki, T. Iwata, T. Ishikawa *et al.*: *Role of Gln1029 in the Photoactivation Processes of the LOV2 Domain in Adiantum Phytochrome3.* Biochemistry **43**, 8373 (2004).

[60] P. L. Freddolino, M. Dittrich, und K. Schulten: *Dynamic Switching Mechanisms in LOV1 and LOV2 Domains of Plant Phototropins*. Biophys. J. **91**, 3630 (2006).

[61] B. D. Zoltowski, C. Schwerdtfeger, J. Widom *et al.*: *Conformational switching in the fungal light sensor Vivid*. Science **316** (2007).

[62] A. Losi und W. Gärtner: *Solving Blue-light Riddles: New Lessons from Flavin-binding LOV Photoreceptors*. Photochem. Photobiol. n/a–n/a (2016).

[63] B. W. Matthews, D. H. Ohlendorf, W. F. Anderson *et al.*: *Structure of the DNA-binding region of lac repressor inferred from its homology with cro repressor*. Proc. Natl. Acad. Sci. USA **79**, 1428 (1982).

[64] R. G. Brennan und B. W. Matthews: *The helix-turn-helix DNA binding motif*. J. Biol. Chem. **264**, 1903 (1989).

[65] E. Peter, B. Dick, und S. A. Baeurle: *Signaling pathway of a photoactivable Rac1-GTPase in the early stages*. Proteins Struct. Funct. Bioinforma. **80**, 1350 (2012).

[66] E. F. Yee, R. P. Diensthuber, A. T. Vaidya *et al.*: *Signal transduction in light–oxygen–voltage receptors lacking the adduct-forming cysteine residue*. Nat. Commun. **6**, 10079 (2015).

[67] J. P. Zayner und T. R. Sosnick: *Factors That Control the Chemistry of the LOV Domain Photocycle*. PLoS One **9**, e87074 (2014).

[68] S. M. Harper, J. M. Christie, und K. H. Gardner: *Disruption of the LOV-Jα Helix Interaction Activates Phototropin Kinase Activity*. Biochemistry **43**, 16184 (2004).

[69] E. Herman, M. Sachse, P. G. Kroth *et al.*: *Blue-Light-Induced Unfolding of the Jα Helix Allows for the Dimerization of Aureochrome-LOV from the Diatom Phaeodactylum tricornutum*. Biochemistry **52**, 3094 (2013).

[70] X. Yao, M. K. Rosen, und K. H. Gardner: *Estimation of the available free energy in a LOV2-J alpha photoswitch*. Nat. Chem Biol. **4** (2008).

[71] S. Usherenko, H. Stibbe, M. Muscó *et al.*: *Photo-sensitive degron variants for tuning protein stability by light*. BMC Syst. Biol. **8**, 128 (2014).

[72] Y. Aihara, T. Yamamoto, K. Okajima *et al.*: *Mutations in N-terminal flanking region of blue light-sensing light-oxygen and voltage 2 (LOV2) domain disrupt its repressive activity on kinase domain in the Chlamydomonas phototropin*. J. Biol. Chem. **287** (2012).

[73] A. I. Nash, W.-H. Ko, S. M. Harper *et al.*: *A Conserved Glutamine Plays a Central Role in LOV Domain Signal Transmission and Its Duration*. Biochemistry **47**, 13842 (2008).

[74] O. I. Lungu, R. A. Hallett, E. J. Choi *et al.*: *Designing photoswitchable peptides using the AsLOV2 domain*. Chem. Biol. **19** (2012).

[75] D. Strickland, K. Moffat, und T. R. Sosnick: *Light-activated DNA binding in a designed allosteric protein.* Proc. Natl. Acad. Sci. USA **105** (2008).

[76] A. Losi und W. Gartner: *Bacterial bilin- and flavin-binding photoreceptors.* Photochem. Photobiol. Sci. **7**, 1168 (2008).

[77] Y. Liu, Q. He, und P. Cheng: *Photoreception in Neurospora: a tale of two White Collar proteins.* Cell Mol. Life Sci. C. **60**, 2131 (2003).

[78] W. R. Briggs und J. M. Christie: *Phototropins 1 and 2: versatile plant blue-light receptors.* Trends Plant Sci. **7**, 204 (2002).

[79] J. M. Christie, M. Salomon, K. Nozue *et al.*: *LOV (light, oxygen, or voltage) domains of the blue-light photoreceptor phototropin (nph1): Binding sites for the chromophore flavin mononucleotide.* Proc. Natl. Acad. Sci. USA **96**, 8779 (1999).

[80] T. Kottke, J. Heberle, D. Hehn *et al.*: *Phot-LOV1: Photocycle of a Blue-Light Receptor Domain from the Green Alga Chlamydomonas reinhardtii.* Biophys. J. **84**, 1192 (2003).

[81] E. Schleicher, R. M. Kowalczyk, C. W. M. Kay *et al.*: *On the Reaction Mechanism of Adduct Formation in LOV Domains of the Plant Blue-Light Receptor Phototropin.* J. Am. Chem. Soc. **126**, 11067 (2004).

[82] K. Lanzl, M. v. Sanden-Flohe, R.-J. Kutta *et al.*: *Photoreaction of mutated LOV photoreceptor domains from Chlamydomonas reinhardtii with aliphatic mercaptans: implications for the mechanism of wild type LOV.* Phys. Chem. Chem. Phys. **12**, 6594 (2010).

[83] V. Dragnea, M. Waegele, S. Balascuta *et al.*: *Time-Resolved Spectroscopic Studies of the AppA Blue-Light Receptor BLUF Domain from Rhodobacter sphaeroides.* Biochemistry **44**, 15978 (2005).

[84] M. Gauden, I. H. M. van Stokkum, J. M. Key *et al.*: *Hydrogen-bond switching through a radical pair mechanism in a flavin-binding photoreceptor.* Proc. Natl. Acad. Sci. USA **103**, 10895 (2006).

[85] K. S. Conrad, C. C. Manahan, und B. R. Crane: *Photochemistry of flavoprotein light sensors.* Nat. Chem Biol. **10**, 801 (2014).

[86] K. Okajima, Y. Fukushima, H. Suzuki *et al.*: *Fate Determination of the Flavin Photoreceptions in the Cyanobacterial Blue Light Receptor TePixD (Tll0078).* J. Mol. Biol. **363**, 10 (2006).

[87] C. Bonetti, M. Stierl, T. Mathes *et al.*: *The Role of Key Amino Acids in the Photoactivation Pathway of the Synechocystis Slr1694 BLUF Domain.* Biochemistry **48**, 11458 (2009).

[88] T. Mathes, I. H. M. van Stokkum, M. Stierl *et al.*: *Redox Modulation of Flavin and Tyrosine Determines Photoinduced Proton-coupled Electron Transfer and Photoactivation of BLUF Photoreceptors.* J. Biol. Chem. **287**, 31725 (2012).

[89] A. A. Gil, A. Haigney, S. P. Laptenok *et al.*: *Mechanism of the AppABLUF Photocycle Probed by Site-Specific Incorporation of Fluorotyrosine Residues: Effect of the Y21 pKa on the Forward and Reverse Ground-State Reactions.* J. Am. Chem. Soc. **138**, 926 (2016).

[90] H. Suzuki, K. Okajima, M. Ikeuchi *et al.*: *LOV-Like Flavin-Cys Adduct Formation by Introducing a Cys Residue in the BLUF Domain of TePixD.* J. Am. Chem. Soc. **130**, 12884 (2008).

[91] C. W. M. Kay, E. Schleicher, A. Kuppig *et al.*: *Blue Light Perception in Plants: Detection and Characterization of a Light-Induced Neutral Flavin Radical in a C450A Mutant of Phototropin.* J. Biol. Chem. **278**, 10973 (2003).

[92] G. Nöll, G. Hauska, P. Hegemann *et al.*: *Redox Properties of LOV Domains: Chemical versus Photochemical Reduction, and Influence on the Photocycle.* ChemBioChem **8**, 2256 (2007).

[93] T. Kottke, J. Heberle, D. Hehn *et al.*: *Phot-LOV1: Photocycle of a Blue-Light Receptor Domain from the Green Alga Chlamydomonas reinhardtii.* Biophys. J. **84**, 1192 (2003).

[94] A. Lukacs, R. Brust, A. Haigney *et al.*: *BLUF Domain Function Does Not Require a Metastable Radical Intermediate State.* J. Am. Chem. Soc. **136**, 4605 (2014).

[95] K. Deisseroth: *Optogenetics.* Nat Meth **8**, 26 (2011).

[96] A. Moglich und K. Moffat: *Engineered photoreceptors as novel optogenetic tools.* Photochem. Photobiol. Sci. **9**, 1286 (2010).

[97] S. Chapman, C. Faulkner, E. Kaiserli *et al.*: *The photoreversible fluorescent protein iLOV outperforms GFP as a reporter of plant virus infection.* Proc. Natl. Acad. Sci. **105**, 20038 (2008).

[98] Y. I. Wu, D. Frey, O. I. Lungu *et al.*: *A genetically encoded photoactivatable Rac controls the motility of living cells.* Nature **461**, 104 (2009).

[99] R. J. Kutta, E. S. A. Hofinger, H. Preuss *et al.*: *Blue-Light Induced Interaction of LOV Domains from Chlamydomonas reinhardtii.* ChemBioChem **9**, 1931 (2008).

[100] R.-J. Kutta, T. Langenbacher, U. Kensy *et al.*: *Setup and performance of a streak camera apparatus for transient absorption measurements in the ns to ms range.* Appl. Phys. B **111**, 203 (2013).

[101] A. Šali und T. L. Blundell: *Comparative Protein Modelling by Satisfaction of Spatial Restraints.* J. Mol. Biol. **234**, 779 (1993).

[102] E. Lindahl, B. Hess, und D. van der Spoel: *GROMACS 3.0: a package for molecular simulation and trajectory analysis.* Mol. Model. Annu. **7**, 306 (2001).

[103] T. A. Soares, X. Daura, C. Oostenbrink *et al.*: *Validation of the GROMOS force-field parameter set 45A3 against nuclear magnetic resonance data of hen egg lysozyme*. J. Biomol. NMR **30**, 407 (2004).

[104] N. Todorova, F. S. Legge, H. Treutlein *et al.*: *Systematic Comparison of Empirical Forcefields for Molecular Dynamic Simulation of Insulin*. J. Phys. Chem. B **112**, 11137 (2008).

[105] D. Frenkel und B. Smit: *Understanding molecular simulation: from algorithms to applications*, Band 1. Academic press (2002).

[106] A. Kotaki, M. Naoi, und K. Yagi: *Effect of Proton Donors on the Absorption Spectrum of Flavin Compounds in Apolar Media*. J. Biochem. **68**, 287 (1970).

[107] S. Salzmann, J. Tatchen, und C. M. Marian: *The photophysics of flavins: What makes the difference between gas phase and aqueous solution?*. J. Photochem. Photobiol. A Chem. **198**, 221 (2008).

[108] M. Sakai und H. Takahashi: *One-electron photoreduction of flavin mononucleotide: time-resolved resonance Raman and absorption study*. J. Mol. Struct. **379**, 9 (1996).

[109] F. Müller, M. Brüstlein, P. Hemmerich *et al.*: *Light-Absorption Studies on Neutral Flavin Radicals*. Eur. J. Biochem. **25**, 573 (1972).

[110] S.-H. Song, B. Dick, A. Penzkofer *et al.*: *Photo-reduction of flavin mononucleotide to semiquinone form in LOV domain mutants of blue-light receptor phot from Chlamydomonas reinhardtii*. J. Photochem. Photobiol. B Biol. **87**, 37 (2007).

[111] D. A. Proshlyakov: *UV optical absorption by protein radicals in cytochrome c oxidase*. Biochim. Biophys. Acta - Bioenerg. **1655**, 282 (2004).

[112] J. Torra, A. Burgos-Caminal, S. Endres *et al.*: *Singlet oxygen photosensitisation by the fluorescent protein Pp2FbFP L30M, a novel derivative of Pseudomonas putida flavin-binding Pp2FbFP*. Photochem. Photobiol. Sci. **14**, 280 (2015).

[113] T. K. Harris und G. J. Turner: *Structural Basis of Perturbed pKa Values of Catalytic Groups in Enzyme Active Sites*. IUBMB Life **53**, 85 (2002).

[114] R. Anandakrishnan, B. Aguilar, und A. V. Onufriev: *H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations*. Nucl. Acids Res. **40**, W537 (2012).

[115] W. Holzer, A. Penzkofer, M. Fuhrmann *et al.*: *Spectroscopic characterization of flavin mononucleotide bound to the LOV1 domain of Phot1 from Chlamydomonas reinhardtii*. Photochem. Photobiol. **75**, 479 (2002).

[116] K. Yagi, N. Ohishi, K. Nishimoto *et al.*: *Effect of hydrogen bonding on electronic spectra and reactivity of flavins*. Biochemistry **19**, 1553 (1980).

[117] *The PyMOL Molecular Graphics System*. PyMOL Molecular Graphics System Version 1.7, Schrödinger, LLC .

[118] J. Zhu, T. Mathes, Y. Hontani *et al.*: *Photoadduct Formation from the FMN Singlet Excited State in the LOV2 Domain of Chlamydomonas reinhardtii Phototropin*. J. Phys. Chem. Lett. 4380–4384 (2016).

[119] T. Biskup, K. Hitomi, E. D. Getzoff *et al.*: *Unexpected Electron Transfer in Cryptochrome Identified by Time-Resolved EPR Spectroscopy*. Angew. Chemie Int. Ed. **50**, 12647 (2011).

[120] G. Lüdemann, I. A. Solov'yov, T. Kubar *et al.*: *Solvent driving force ensures fast formation of a persistent and well-separated radical pair in plant cryptochrome*. J. Am. Chem. Soc. **137**, 1147 (2015).

[121] J. J. Warren, J. R. Winkler, und H. B. Gray: *Redox properties of tyrosine and related molecules*. FEBS Lett. **586**, 596 (2012).

[122] S. Y. Reece und D. G. Nocera: *Proton-coupled electron transfer in biology: results from synergistic studies in natural and model systems*. Annu. Rev. Biochem. **78**, 673 (2009).

[123] I. A. Solov'yov, T. Domratcheva, A. R. M. Shahi *et al.*: *Decrypting cryptochrome: Revealing the molecular identity of the photoactivation reaction*. J. Am. Chem. Soc. **134**, 18046 (2012).

[124] I. A. Solov'yov, T. Domratcheva, und K. Schulten: *Separation of photo-induced radical pair in cryptochrome to a functionally critical distance*. Sci. Rep. **4**, 3845 (2014).

[125] T. Langenbacher, D. Immeln, B. Dick *et al.*: *Microsecond light-induced proton transfer to flavin in the blue light sensor plant cryptochrome*. J. Am. Chem. Soc. **131**, 14274 (2009).

[126] C. Thöing, S. Oldemeyer, und T. Kottke: *Microsecond Deprotonation of Aspartic Acid and Response of the $\alpha/\beta$ Subdomain Precede C-Terminal Signaling in the Blue Light Sensor Plant Cryptochrome*. J. Am. Chem. Soc. **137**, 5990 (2015).

[127] P. Müller, J.-P. Bouly, K. Hitomi *et al.*: *ATP Binding Turns Plant Cryptochrome Into an Efficient Natural Photoswitch*. Sci. Rep. **4**, 5175 (2014).

[128] R. D. Draper und L. L. Ingraham: *A Potentiometric Study of the Flavin Semiquinone Equilibrium*. Arch. Biochem. Biophys. **125**, 802 (1968).

[129] H. Pelletier und J. Kraut: *Crystal structure of a complex between electron transfer partners, cytochrome c peroxidase and cytochrome c*. Science **258**, 1748 (1992).

[130] P. Bertrand: *Application of electron transfer theories to biological systems*. In *Long-Range Electron Transf. Biol.*, 1–47. Springer Berlin Heidelberg, Berlin, Heidelberg(1991).

[131] H. B. Gray und J. R. Winkler: *Long-range electron transfer*. Proc. Natl. Acad. Sci. USA **102**, 3534 (2005).

[132] L. A. Abriata, D. Álvarez-Paggi, G. N. Ledesma *et al.*: *Alternative ground states enable pathway switching in biological electron transfer*. Proc. Natl. Acad. Sci. USA **109**, 17348 (2012).

[133] R. A. Marcus und N. Sutin: *Electron transfers in chemistry and biology*. Biochim. Biophys. Acta - Rev. Bioenerg. **811**, 265 (1985).

[134] V. G. Levich und R. R. Dogonadze: *Theory of non-radiation electron transitions from ion to ion in solutions*. Dokl. Akad. Nauk SSSR **124**, 123 (1959).

[135] N. S. Hush: *Homogeneous and heterogeneous optical and thermal electron transfer*. Electrochim. Acta **13**, 1005 (1968).

[136] M. R. Gunner und P. L. Dutton: *Temperature and $\Delta G°$ dependence of the electron transfer from $BPh^{\cdot-}$ to QA in reaction center protein from Rhodobacter sphaeroides with different quinones as QA*. J. Am. Chem. Soc. **111**, 3400 (1989).

[137] G. Venturoli, F. Drepper, J. C. Williams *et al.*: *Effects of Temperature and $\Delta G°$ on Electron Transfer from Cytochrome c2 to the Photosynthetic Reaction Center of the Purple Bacterium Rhodobacter sphaeroides*. Biophys. J. **74**, 3226 (1998).

[138] A. R. Crofts und S. Rose: *Marcus treatment of endergonic reactions: A commentary*. Biochim. Biophys. Acta - Bioenerg. **1767**, 1228 (2007).

[139] C. C. Page, C. C. Moser, und P. L. Dutton: *Mechanism for electron transfer within and between proteins*. Curr. Opin. Chem. Biol. **7**, 551 (2003).

[140] D. N. Beratan und J. J. Hopfield: *Calculation of tunneling matrix elements in rigid systems: mixed-valence dithiaspirocyclobutane molecules*. J. Am. Chem. Soc. **106**, 1584 (1984).

[141] A. Kuki und P. G. Wolynes: *Electron tunneling paths in proteins*. Science **236**, 1647 (1987).

[142] I. V. Kurnikov: *HARLEM molecular modeling package*. Dep. Chem. Univ. Pittsburgh, Pittsburgh, PA (2000).

[143] J. Blumberger: *Free energies for biological electron transfer from QM/MM calculation: method, application and critical assessment*. Phys. Chem. Chem. Phys. **10**, 5651 (2008).

[144] V. Guallar und F. Wallrapp: *Mapping protein electron transfer pathways with QM/MM methods*. J. R. Soc. Interface **5**, 233 (2008).

[145] E. Brunk und U. Rothlisberger: *Mixed Quantum Mechanical/Molecular Mechanical Molecular Dynamics Simulations of Biological Systems in Ground and Electronically Excited States*. Chem. Rev. **115**, 6217 (2015).

[146] J. Novotny, R. E. Bruccoleri, und F. A. Saul: *On the attribution of binding energy in antigen-antibody complexes McPC 603, D1.3, and HyHEL-5*. Biochemistry **28**, 4735 (1989).

[147] J. Novotny, R. E. Bruccoleri, M. Davis *et al.*: *Empirical free energy calculations: a blind test and further improvements to the method 1*. J. Mol. Biol. **268**, 401 (1997).

[148] S. Y. Reece, J. M. Hodgkiss, J. Stubbe *et al.*: *Proton-coupled electron transfer: the mechanistic underpinning for radical transport and catalysis in biology*. Philos. Trans. R. Soc. B Biol. Sci. **361**, 1351 (2006).

[149] A. Lukacs, A. P. M. Eker, M. Byrdin *et al.*: *Electron Hopping through the 15 Å Triple Tryptophan Molecular Wire in DNA Photolyase Occurs within 30 ps*. J. Am. Chem. Soc. **130**, 14394 (2008).

[150] D. Immeln, A. Weigel, T. Kottke *et al.*: *Primary Events in the Blue Light Sensor Plant Cryptochrome: Intraprotein Electron and Proton Transfer Revealed by Femtosecond Spectroscopy*. J. Am. Chem. Soc. **134**, 12536 (2012).

[151] A. de la Lande, N. S. Babcock, J. Rezáč *et al.*: *Surface residues dynamically organize water bridges to enhance electron transfer between proteins*. Proc. Natl. Acad. Sci. USA **107**, 11799 (2010).

[152] C. Shih, A. K. Museth, M. Abrahamsson *et al.*: *Tryptophan-Accelerated Electron Flow Through Proteins*. Science **320**, 1760 (2008).

[153] I. F. Sevrioukova, H. Li, H. Zhang *et al.*: *Structure of a cytochrome P450–redox partner electron-transfer complex*. Proc. Natl. Acad. Sci. USA **96**, 1863 (1999).

[154] C. Consani, G. Auböck, F. van Mourik *et al.*: *Ultrafast Tryptophan-to-Heme Electron Transfer in Myoglobins Revealed by UV 2D Spectroscopy*. Science **339**, 1586 (2013).

[155] M. G. Khrenova, A. V. Nemukhin, und T. Domratcheva: *Photoinduced Electron Transfer Facilitates Tautomerization of the Conserved Signaling Glutamine Side Chain in BLUF Protein Light Sensors*. J. Phys. Chem. B **117**, 2369 (2013).

[156] H. W. Park, S. T. Kim, A. Sancar *et al.*: *Crystal structure of DNA photolyase from Escherichia coli*. Science **268**, 1866 (1995).

[157] Z. Chen, R. Durley, V. Davidson *et al.*: *Structural comparison of the oxidized ternary electron transfer complex of methylamine dehydrogenase, amicyanin and cytochrome c551i from Paracoccus denitrificans with the substrate-reduced, copper free complex at 1.9 Å resolution*. TBP DOI: 10.2210/pdb2gc4/pdb (2006).

[158] A. Blanco-Rodriguez, M. Busby, K. Ronayne *et al.*: *Relaxation dynamics of Pseudomonas aeruginosa Re(I)(CO)3(alpha-diimine)(HisX)+ (X = 83, 107, 109, 124, 126)Cu(II) azurins*. J. Am. Chem. Soc. **131**, 11788 (2009).

[159] M. Hogbom, M. Galander, M. Andersson *et al.*: *Displacement of the tyrosyl radical cofactor in ribonucleotide reductase obtained by single-crystal high-field EPR and 1.4-A x-ray data.* Proc. Natl. Acad. Sci. USA **100**, 3209 (2003).

[160] S. Evans und G. Brayer: *High-resolution study of the three-dimensional structure of horse heart metmyoglobin.* J. Mol. Biol. **213**, 885 (1990).

[161] A. Jung, J. Reinstein, T. Domratcheva *et al.*: *Crystal Structures of the AppA BLUF Domain Photoreceptor Provide Insights into Blue Light-mediated Signal Transduction.* J. Mol. Biol. **362**, 717 (2006).

[162] C. Brautigam, B. Smith, Z. Ma *et al.*: *Structure of the photolyase-like domain of cryptochrome 1 from Arabidopsis thaliana.* Proc. Natl. Acad. Sci. USA **101**, 12142 (2004).

[163] M. Zhang, L. Wang, S. Shu *et al.*: *Bifurcating electron-transfer pathways in DNA photolyases determine the repair quantum yield.* Science **354**, 209 (2016).

[164] I. M. M. Wijaya, T. Domratcheva, T. Iwata *et al.*: *Single Hydrogen Bond Donation from Flavin N5 to Proximal Asparagine Ensures FAD Reduction in DNA Photolyase.* J. Am. Chem. Soc. **138**, 4368 (2016).

[165] D. Zhong: *Electron Transfer Mechanisms of DNA Repair by Photolyase.* Annu. Rev. Phys. Chem. **66**, 691 (2015).

[166] P. Scheerer, F. Zhang, J. Kalms *et al.*: *The Class III Cyclobutane Pyrimidine Dimer Photolyase Structure Reveals a New Antenna Chromophore Binding Site and Alternative Photoreduction Pathways.* J. Biol. Chem. **290**, 11504 (2015).

[167] M. Biasini, S. Bienert, A. Waterhouse *et al.*: *SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information.* Nucl. Acids Res. **42**, W252 (2014).

[168] L. Orlichenko, R. Geyer, M. Yanagisawa *et al.*: *The 19-amino acid insertion in the tumor-associated splice isoform Rac1b confers specific binding to p120 catenin.* J. Biol. Chem. **285**, 19153 (2010).

[169] A. B. Jaffe und A. Hall: *Rho GTPases: Biochemistry and Biology.* Annu. Rev. Cell Dev. Biol. **21**, 247 (2005).

[170] A. J. Ridley: *Rho GTPases and actin dynamics in membrane protrusions and vesicle trafficking.* Trends Cell Biol. **16**, 522 (2006).

[171] P. Aspenström, Å. Fransson, und J. Saras: *Rho GTPases have diverse effects on the organization of the actin filament system.* Biochem. J. **377**, 327 LP (2004).

[172] C. J. D. Irene M Zohn, Sharon L Campbell, Roya Khosravi-Far, Kent L Rossman: *Rho family proteins and Ras transformation: the RHOad less traveled gets congested.* Oncogene **17**, 1415 (1998).

[173] D. K. Worthylake, K. L. Rossman, und J. Sondek: *Crystal structure of Rac1 in complex with the guanine nucleotide exchange region of Tiam1*. Nature **408**, 682 (2000).

[174] A. Schnelzer, D. Prechtel, U. Knaus *et al.*: *Rac1 in human breast cancer: overexpression, mutation analysis, and characterization of a new isoform, Rac1b*. Oncogene **19**, 3013 (2000).

[175] D. Fiegen, L.-C. Haeusler, L. Blumenstein *et al.*: *Alternative splicing of Rac1 generates Rac1b, a self-activating GTPase*. J. Biol. Chem. **279**, 4743 (2004).

[176] K. Wennerberg, K. L. Rossman, und C. J. Der: *The Ras superfamily at a glance*. J. Cell Sci. **118**, 843 LP (2005).

[177] M. Parri und P. Chiarugi: *Rac and Rho GTPases in cancer cell motility control*. Cell Commun. Signal. **8**, 1 (2010).

[178] Y. Gu, M.-D. Filippi, J. A. Cancelas *et al.*: *Hematopoietic Cell Regulation by Rac1 and Rac2 Guanosine Triphosphatases*. Science **302**, 445 LP (2003).

[179] L. Menard, E. Tomhave, P. J. Casey *et al.*: *Rac1, a low-molecular-mass GTP-binding-protein with high intrinsic GTPase activity and distinct biochemical properties*. Eur. J. Biochem. **206**, 537 (1992).

[180] M. Krauthammer, Y. Kong, B. H. Ha *et al.*: *Exome sequencing identifies recurrent somatic RAC1 mutations in melanoma*. Nat. Genet. **44**, 1006 (2012).

[181] N. Sawada, Y. Li, und J. K. Liao: *Novel aspects of the roles of Rac1 GTPase in the cardiovascular system*. Curr. Opin. Pharmacol. **10**, 116 (2010).

[182] J. L. MacKay und S. Kumar: *Simultaneous and independent tuning of RhoA and Rac1 activity with orthogonally inducible promoters*. Integr. Biol. **6**, 885 (2014).

[183] S. Etienne-Manneville und A. Hall: *Rho GTPases in cell biology*. Nature **420** (2002).

[184] X. R. Bustelo, V. Sauzeau, und I. M. Berenjeno: *GTP-binding proteins of the Rho/Rac family: regulation, effectors and functions in vivo*. Bioessays **29**, 356 (2007).

[185] S. J. Heasman und A. J. Ridley: *Mammalian Rho GTPases: new insights into their functions from in vivo studies*. Nat. Rev. Mol. Cell Biol. **9**, 690 (2008).

[186] H. Kim und J. K. Sonn: *Rac1 promotes chondrogenesis by regulating STAT3 signaling pathway*. Cell Biol. Int. **40**, 976 (2016).

[187] H. Teramoto, O. A. Coso, H. Miyata *et al.*: *Signaling from the Small GTP-binding Proteins Rac1 and Cdc42 to the c-Jun N-terminal Kinase/Stress-activated Protein Kinase Pathway: A Role for Mixed Lineage Kinase 3/Protein-Tyrosine Kinase 1, a Novel Member of the Mixed Lineage Kinase Family*. J. Biol. Chem. 27225–27228 .

[188] M. Nagase und T. Fujita: *Role of Rac1-mineralocorticoid-receptor signalling in renal and cardiac disease*. Nat. Rev. Nephrol. **9**, 86 (2013).

[189] G. Fritz, I. Just, und B. Kaina: *Rho GTPases are over-expressed in human tumors*. Int. J. Cancer **81** (1999).

[190] C. Zhou, S. Licciulli, J. L. Avila *et al.*: *The Rac1 splice form Rac1b promotes K-ras-induced lung tumorigenesis*. Oncogene **32**, 903 (2013).

[191] M. L. Stallings-Mann, J. Waldmann, Y. Zhang *et al.*: *Matrix Metalloproteinase Induction of Rac1b, a Key Effector of Lung Cancer Progression*. Sci. Transl. Med. **4**, 142ra95 LP (2012).

[192] A. L. Silva, F. Carmo, und M. J. Bugalho: *Rac1b overexpression in papillary thyroid carcinoma: a role to unravel*. Eur. J. Endocrinol. **168**, 795 (2013).

[193] Márcia Faria, Liliana Capinha, Joana Simões-Pereira, Maria João Bugalho und A. L. Silva: *Extending the Impact of Rac1b Overexpression to Follicular Thyroid Carcinomas*. Int. J. Endocrinol. **2016**, 1 (2016).

[194] C. Mehner, E. Miller, D. Khauv *et al.*: *Tumor Cell–Derived MMP3 Orchestrates Rac1b and Tissue Alterations That Promote Pancreatic Adenocarcinoma*. Mol. Cancer Res. **12**, 1430 LP (2014).

[195] P. Matos, J. G. Collard, und P. Jordan: *Tumor-related alternatively spliced Rac1b is not regulated by Rho-GDP dissociation inhibitors and exhibits selective downstream signaling*. J. Biol. Chem. **278** (2003).

[196] C. G. Peter Jordan, Raquel Braza, Maria Guida Boavida und E. Chastre: *Cloning of a novel human Rac1b splice variant with increased expression in colorectal tumors*. Oncogene **18**, 6835 (1999).

[197] P. Matos und P. Jordan: *Rac1, but Not Rac1B, Stimulates RelB-mediated Gene Transcription in Colorectal Cancer Cells*. J. Biol. Chem. **281**, 13724 (2006).

[198] G. Li, L. Ying, H. Wang *et al.*: *Rac1b enhances cell survival through activation of the JNK2/c-JUN/Cyclin-D1 and AKT2/MCL1 pathways*. Oncotarget **7**, 17970 (2016).

[199] K. Bedard und K.-H. Krause: *The NOX Family of ROS-Generating NADPH Oxidases: Physiology and Pathophysiology*. Physiol. Rev. **87**, 245 LP (2007).

[200] V. Braunersreuther, F. Montecucco, M. Ashri *et al.*: *Role of NADPH oxidase isoforms NOX1, NOX2 and NOX4 in myocardial ischemia/reperfusion injury*. J. Mol. Cell Cardiol. **64**, 99 (2013).

[201] K. Lapouge, S. J. Smith, P. A. Walker *et al.*: *Structure of the TPR Domain of p67phox in Complex with Rac · GTP*. Mol. Cell **6**, 899 (2000).

[202] P. M.-C. Dang, H. Raad, R. A. Derkawi *et al.*: *The NADPH oxidase cytosolic component p67phox is constitutively phosphorylated in human neutrophils:*

*Regulation by a protein tyrosine kinase, MEK1/2 and phosphatases 1/2A.* Biochem. Pharmacol. **82**, 1145 (2011).

[203] C.-H. Han, J. L. R. Freeman, T. Lee *et al.*: *Regulation of the Neutrophil Respiratory Burst Oxidase: Identification of an Activation Domain in p67phox.* J. Biol. Chem. **273**, 16663 (1998).

[204] A. Dovas und J. R. Couchman: *RhoGDI: multiple functions in the regulation of Rho family GTPase activities.* Biochem. J. **390**, 1 LP (2005).

[205] T. Hakoshima, T. Shimizu, und R. Maesaki: *Structural Basis of the Rho GTPase Signaling.* J. Biochem. **134**, 327 (2003).

[206] A. Winkler, T. R. M. Barends, A. Udvarhelyi *et al.*: *Structural Details of Light Activation of the LOV2-based Photoswitch PA-Rac1.* ACS Chem. Biol. **10**, 502 (2015).

[207] S. Grizot, J. Fauré, F. Fieschi *et al.*: *Crystal Structure of the Rac1-RhoGDI Complex Involved in NADPH Oxidase Activation,.* Biochemistry **40**, 10007 (2001).

[208] M. V. Milburn, L. Tong, A. M. DeVos *et al.*: *Molecular switch for signal transduction: structural differences between active and inactive forms of protooncogenic ras proteins.* Science **247**, 939 LP (1990).

[209] J. John, H. Rensland, I. Schlichting *et al.*: *Kinetic and structural analysis of the Mg(2+)-binding site of the guanine nucleotide-binding protein p21H-ras.* J. Biol. Chem. **268**, 923 (1993).

[210] J. Y. Pan und M. Wessling-Resnick: *GEF-mediated GDP/GTP exchange by monomeric GTPases: A regulatory role for Mg2+?.* Bioessays **20**, 516 (1998).

[211] A. Kumar, V. Rajendran, R. Sethumadhavan *et al.*: *Molecular Dynamic Simulation Reveals Damaging Impact of RAC1 F28L Mutation in the Switch I Region.* PLoS One **8**, e77453 (2013).

[212] Y. Ugolev, Y. Berdichevsky, C. Weinbaum *et al.*: *Dissociation of Rac1(GDP) · RhoGDI Complexes by the Cooperative Action of Anionic Liposomes Containing Phosphatidylinositol 3,4,5-Trisphosphate, Rac Guanine Nucleotide Exchange Factor, and GTP.* J. Biol. Chem. **283**, 22257 (2008).

[213] T. Ueyama, J. Son, T. Kobayashi *et al.*: *Negative Charges in the Flexible N-Terminal Domain of Rho GDP-Dissociation Inhibitors (RhoGDIs) Regulate the Targeting of the RhoGDI–Rac1 Complex to Membranes.* J. Immunol. **191**, 2560 (2013).

[214] M. T. A. Alexandre, T. Domratcheva, C. Bonetti *et al.*: *Primary Reactions of the LOV2 Domain of Phototropin Studied with Ultrafast Mid-Infrared Spectroscopy and Quantum Chemistry.* Biophys. J. **97**, 227 (2009).

[215] E. Peter, B. Dick, und S. A. Baeurle: *Regulatory mechanism of the light-activable allosteric switch LOV–TAP for the control of DNA binding: A computer simulation study.* Proteins Struct. Funct. Bioinforma. **81**, 394 (2012).

[216] M. A. Martí-Renom, A. C. Stuart, A. Fiser *et al.*: *Comparative Protein Structure Modeling of Genes and Genomes.* Annu. Rev. Biophys. Biomol. Struct. **29**, 291 (2000).

[217] C. Chothia und A. M. Lesk: *The relation between the divergence of sequence and structure in proteins.* EMBO J. **5**, 823 (1986).

[218] M. Lingenheil, R. Denschlag, R. Reichold *et al.*: *The "Hot-Solvent/Cold-Solute" Problem Revisited.* J. Chem. Theory Comput. **4**, 1293 (2008).

[219] B. Hess, H. Bekker, H. J. C. Berendsen *et al.*: *LINCS: A linear constraint solver for molecular simulations.* J. Comput. Chem. **18**, 1463 (1997).

[220] C. Neiss und P. Saalfrank: *Molecular Dynamics Simulation of the LOV2 Domain from Adiantum capillus-veneris.* J. Chem. Inf. Comput. Sci. **44**, 1788 (2004).

[221] N. A. Baker, D. Sept, S. Joseph *et al.*: *Electrostatics of nanosystems: Application to microtubules and the ribosome.* Proc. Natl. Acad. Sci. USA **98**, 10037 (2001).

[222] M. Lerner und H. Harlson: *APBS plugin for PyMOL.* Ann. Arbor Univ. Michigan (2006).

[223] T. J. Dolinsky, J. E. Nielsen, J. A. McCammon *et al.*: *PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations.* Nucl. Acids Res. **32**, W665 (2004).

[224] T. J. Dolinsky, P. Czodrowski, H. Li *et al.*: *PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations.* Nucl. Acids Res. **35**, W522 (2007).

[225] A. D. MacKerell, N. Banavali, und N. Foloppe: *Development and current status of the CHARMM force field for nucleic acids.* Biopolymers **56**, 257 (2000).

[226] A. W. Schüttelkopf und D. M. F. van Aalten: *PRODRG: a tool for high-throughput crystallography of protein-ligand complexes.* Acta Crystallogr. Sect. D **60**, 1355 (2004).

[227] W. Kabsch und C. Sander: *Dictionary of Protein Secondary Structure - Pattern-Recognition of Hydrogen-Bonded and Geometrical Features.* Biopolymers **22** (1983).

[228] W. G. Touw, C. Baakman, J. Black *et al.*: *A series of PDB-related databanks for everyday needs.* Nucl. Acids Res. **43**, D364 (2014).